UNIVERSITY OF LIVERPOOL

Measurement of the top quark pair production cross-section in the μeb final state at $\sqrt{s}$=13 TeV with the LHCb detector.

by

Heather McKenzie Wark

Department of Physics
Oliver Lodge Laboratory
University of Liverpool

A thesis submitted in partial fulfilment for the degree of Doctor of Philosophy

March 2020

Declaration of Authorship

This thesis is a result of my own work, except where a specific reference to the work of others is made. This thesis has not been submitted for any other qualification to this or any other university.

Heather McKenzie Wark

'Around here, however, we don't look backwards for very long.
We ***keep moving forward****, opening up new doors*
and doing new things because we're curious...
and curiosity keeps leading us down new paths.'

Walt Disney

Abstract

The LHC is the highest energy particle collider in the world. The LHCb experiment is one of the four main experiments at the LHC. Between 2015 and 2018 it collected 5.4 fb^{-1} of data at a centre of mass energy of 13 TeV. In this thesis, the cross-section for $t\bar{t} \to \mu eb$ production is measured using this data set. The study of this measurement is a test of the predicting power of Standard Model theory and also contributes to constraining future predictions of proton interactions.

A selection scheme is developed to reduce the expected contributions from a number of backgrounds and improve the purity of the final state. The final measurement, $\sigma_{t\bar{t}}$, represents the cross-section of the production of top quark pairs decaying to a muon, electron and b-jet, where the p_{T} of each object is above 20 GeV, the lepton pseudorapidities are between 2 and 4.5, and the b-jet pseudorapidity is between 2.2 and 4.2. The result is:

$$\sigma_{t\bar{t}} = 117 \pm 10(\text{stat.}) \pm 15(\text{syst.}) \pm 5(\text{lumi.})\text{pb} \quad (1)$$

where the first uncertainty is statistical, the second is systematic and the third is due to the luminosity determination. Comparisons are made with theoretical predictions calculated from POWHEG and aMC@NLO and the measurement is seen to be in good agreement.

A feasibility study of the measurement of the WW production cross-section is also performed. While the measurement will be possible at LHCb in the near future, it is not feasible with the current data set, due mainly to the limited statistics available. It is predicted that a significant measurement of this cross-section will be possible by the end of LHCb Run-III.

Acknowledgements

First and foremost I would like to extend my sincere gratitude to my supervisor Tara Shears and to Stephen Farry for their guidance and patience over the last four years. I would also like to thank the rest of the Liverpool LHCb group for their support and advice, particularly Vin Franco who has gone through each step of this journey with me and has always been happy to offer his (often questionable) physics knowledge.

I could not have made it to the end of this process without the support of some truly fantastic women who have encouraged me to continue through some of the most difficult obstacles I have faced in the last four years. I'd like to thank Helen Vaughan for that poignant cup of tea, Emily Graham for her wonderful common-sense filtering, Lauren Anthony for very late nights in the office, and Tabitha Halewood-Leagas for agreeing we'd make it to the end for each other.

I would also like to thank Barry King, to whose memory this thesis is dedicated and without whom I wouldn't be here today. Barry first sparked my interest in Particle Physics during Sixth Form when I was one of his Nuffield Bursary students. He taught me during my undergraduate degree, he was my Masters project supervisor, he offered so much support and encouragement through the first few years of my postgraduate study. I miss him dearly.

Finally, I would like to thank my family for their love and support, and for always encouraging me to believe I can do anything I set my mind to.

I would like to dedicate this thesis to
the memory of Dr. Barry King.

Contents

Declaration of Authorship iii

Abstract vii

Acknowledgements ix

1 Introduction 1

2 Theoretical Overview 3
2.1 The Standard Model . . . 3
2.1.1 Particles and Forces . . . 3
2.1.1.1 Fermions . . . 4
2.1.1.2 Gauge Bosons . . . 5
2.1.2 Quantum Electrodynamics . . . 5
2.1.3 Quantum Chromodynamics . . . 7
2.1.4 Electroweak Theory . . . 8
2.1.5 The Higgs Mechanism . . . 9
2.1.6 Yukawa Coupling . . . 12
2.1.7 Quark Mixing and CP Violation . . . 13
2.1.8 Summary . . . 14
2.2 Testing the Standard Model at Hadron Colliders . . . 15
2.2.1 Hard Scattering and the Underlying Event . . . 16
2.2.2 Parton Distribution Functions . . . 17
2.3 The Top Quark . . . 18
2.3.1 Motivation . . . 18
2.3.2 Production of Top Quark Pairs . . . 20
2.3.3 Top Quark Decay . . . 21

3 Experimental Environment 23
3.1 The Large Hadron Collider . . . 23
3.1.1 The Accelerator Complex . . . 23
3.1.2 Running Performance . . . 25
3.2 The LHCb Experiment . . . 28
3.3 Tracking System . . . 30
3.3.1 Vertex Locator . . . 30
3.3.2 Magnet . . . 33
3.3.3 Tracker Turicensis . . . 34

3.3.4 Downstream Trackers . . . 34
3.4 RICH . . . 36
3.5 Calorimeter System . . . 37
3.6 Muon System . . . 40

4 Event Reconstruction **43**
4.1 Track Reconstruction . . . 43
4.2 Vertexing . . . 46
4.3 Particle Identification . . . 47
4.3.1 Photon Identification . . . 48
4.3.2 Electron Identification . . . 49
4.3.3 Muon Identification . . . 49
4.3.4 RICH Identification . . . 50
4.3.5 Jet Reconstruction . . . 50
4.4 Triggers . . . 51
4.4.1 L0 . . . 52
4.4.1.1 Calorimeter Trigger . . . 52
4.4.1.2 Muon Trigger . . . 52
4.4.1.3 L0 Decision . . . 53
4.4.2 HLT1 . . . 53
4.4.3 HLT2 . . . 54
4.4.4 Real-time Alignment and Calibration . . . 54
4.5 Software . . . 55

5 Top Quark Pair Production Cross-section **57**
5.1 Motivation . . . 57
5.2 Data and Simulation . . . 61
5.3 Event Selection . . . 62
5.3.1 Trigger Requirements . . . 62
5.3.2 Lepton Selection . . . 62
5.3.2.1 Impact Parameter . . . 62
5.3.2.2 Isolation . . . 63
5.3.2.3 Identification Requirements . . . 64
5.3.3 Jet Selection . . . 66
5.3.4 Summary . . . 66
5.4 Signal Purity . . . 67
5.4.1 Lepton Misidentification . . . 67
5.4.2 Z + jet . . . 68
5.4.3 Single Top . . . 70
5.4.4 Di-boson . . . 71
5.4.5 Misidentified Jets . . . 71
5.4.6 Summary . . . 71
5.5 Efficiencies . . . 72
5.5.1 Reconstruction Efficiencies . . . 73
5.5.1.1 Trigger Efficiency . . . 75
5.5.1.2 Muon Identification Efficiency . . . 77
5.5.1.3 Muon Tracking Efficiency . . . 78

5.5.1.4 Electron Identification Efficiency . . . 79
5.5.1.5 Electron Tracking Efficiency . . . 81
5.5.1.6 Jet Reconstruction Efficiency . . . 85
5.5.2 Jet Tagging Efficiency . . . 87
5.5.3 Selection Efficiency . . . 91
5.5.4 Summary . . . 94
5.6 Resolution Factor . . . 94
5.6.1 Jet Energy Scale . . . 95
5.6.2 Electron Momentum Scale . . . 97
5.6.3 Summary . . . 99
5.7 Cross-section Calculation . . . 99
5.8 Systematic Uncertainties . . . 101
5.9 Results . . . 101

6 Feasibility of a WW production cross-section measurement 105
6.1 Motivation . . . 105
6.2 Data and Simulation . . . 107
6.3 Event Selection . . . 107
6.3.1 Isolation . . . 108
6.3.2 Impact Parameter . . . 108
6.3.3 Jet Multiplicity . . . 108
6.3.4 p_{T} Vector Sum . . . 109
6.3.5 Summary . . . 110
6.4 Signal Purity . . . 112
6.5 Results . . . 113
6.6 Outlook . . . 114

7 Conclusion 119

List of Figures 120

List of Tables 125

Chapter 1

Introduction

The Standard Model (SM) of particle physics is the current theory describing the fundamental constituents of matter and the forces that govern their behaviour. The SM must be tested to validate the precision of the theoretical predictions it makes in order to continue searching for physics beyond the model to better describe our universe.

SM predictions can be tested using particle accelerators. The Large Hadron Collider (LHC) is currently the highest energy particle accelerator in the world, accelerating protons to close to the speed of light before colliding them to produce different particles for study.

In this thesis, the top quark (the heaviest fundamental particle) is investigated through the partial reconstruction of decays of top quark pairs to a muon, electron and b-jet. The frequency with which top quark pairs are produced at the LHC (or the top quark pair production cross-section) is measured using data collected during the second Run of the LHCb experiment between 2015 and 2018. This is an important SM measurement which can be used to further improve our understanding of the proton and so constrain future theoretical predictions of proton interactions.

This thesis is outlined as follows. An introduction to the Standard Model is presented in Chapter 2, followed by a summary of how the SM is tested at particle colliders and a brief discussion of top quark theory. Chapter 3 describes the experimental environment of the LHC and the LHCb detector, and Chapter 4 details the processes required to reconstruct collision events from detector signals for physics analysis.

The primary analysis presented here is a measurement of the top quark pair production cross-section, which is outlined in Chapter 5. A study of the feasibility of measuring the WW production cross-section is also presented in Chapter 6. The conclusions from both analyses are reviewed in Chapter 7.

Chapter 2

Theoretical Overview

This chapter gives a basic overview of the leading theory of particle physics: The Standard Model (SM). This is followed by a discussion of how the Standard Model can be tested at particle colliders such as the LHC, and then by a description of key aspects of SM physics that relate to the analyses presented in later chapters.

2.1 The Standard Model

The Standard Model of Particle Physics is the current theory used to describe the fundamental constituents of matter and the forces through which they interact. The Standard Model is described below, starting with the particles and forces that make up the model, followed by the mathematical framework used to describe their interactions. It is a relativistic Quantum Field Theory (QFT), treating fundamental particles as excited states of quantum fields defined at all points in spacetime.

2.1.1 Particles and Forces

The Standard Model describes the fundamental particles identified in nature (fermions[1]). It also describes three fundamental forces and the particles that mediate these interactions. The theory also includes the Higgs mechanism which is responsible for providing the fundamental particles with mass.

[1] Fermions are defined as particles with half-integer intrinsic spin.

2.1.1.1 Fermions

There are two categories of fermions in the Standard Model: quarks and leptons. There are six of each, categorised into three *generations*.

Each generation of quarks contains two particles separated by one unit of electrical charge. The generational pairs are: the up and down quarks (u and d), the charm and strange quarks (c and s), and the top and bottom quarks (t and b). The up-type quarks have a charge of $+\frac{2}{3}$ and the down-type quarks have $-\frac{1}{3}$. The properties of the quarks are summarised in Table 2.1[2].

Generation	Quark	Mass (MeV)	Spin	Q/e
1	u	$2.16^{+0.49}_{-0.26}$	$\frac{1}{2}$	$+\frac{2}{3}$
	d	$4.67^{+0.48}_{-0.17}$	$\frac{1}{2}$	$-\frac{1}{3}$
2	c	1,270±20	$\frac{1}{2}$	$+\frac{2}{3}$
	s	93^{+11}_{-5}	$\frac{1}{2}$	$-\frac{1}{3}$
3	t	172,900±400	$\frac{1}{2}$	$+\frac{2}{3}$
	b	4180^{+30}_{20}	$\frac{1}{2}$	$-\frac{1}{3}$

TABLE 2.1: The properties of the quarks in the Standard Model [1].

The corresponding generations of leptons each contain a charged particle and its neutral partner. The three charged leptons are the electron (e^-), the muon (μ^-), and the tau (τ^-), each displaying identical characteristics except for their mass and lepton flavour. The neutral leptons have negligible mass and occur in three matching flavours: the electron neutrino (ν_e), the muon neutrino (ν_μ) and the tau neutrino (ν_τ). The properties of the leptons are summarised in Table 2.2.

Generation	Lepton	Mass (MeV)	Spin	Q/e
1	e^-	0.5109989461±0.0000000031	$\frac{1}{2}$	-1
	ν_e	<0.000002	$\frac{1}{2}$	0
2	μ^-	105.6583748±0.0000024	$\frac{1}{2}$	-1
	ν_μ	<0.19	$\frac{1}{2}$	0
3	τ^-	1776.86±0.12	$\frac{1}{2}$	-1
	ν_τ	<18.2	$\frac{1}{2}$	0

TABLE 2.2: The properties of the leptons in the Standard Model [1].

For every fermion there is a corresponding particle with the same mass and opposite-sign quantum numbers. These particles are referred to as *antiparticles* or *antimatter*.

[2]Units of c=$\hbar$=1 are used from this point forward

2.1.1.2 Gauge Bosons

Bosons are spin-1 particles that mediate the fundamental forces: the photon (γ) for the electromagnetic force; $W^{\pm}$ and Z^0 bosons for the weak interaction; and the gluon for the strong. The properties of the bosons are summarised in Table 2.3.

Field	Boson	Mass (GeV)	Spin	Q/e
Electromagnetic	γ	0	1	0
Weak nuclear	$W^{\pm}$	80.379±0.012	1	±1
	Z^0	91.1876±0.0021	1	0
Strong nuclear	g	0	1	0

TABLE 2.3: The properties of the bosons in the Standard Model [1].

2.1.2 Quantum Electrodynamics

Quantum Electrodynamics (QED) is the QFT used to describe electromagnetic interactions. A free Dirac field, $\Psi(x)$, describes the behaviour of a fermion with charge Q and mass m. Its dynamics are described by the Dirac equation [2]

$$(i \not{\partial} - m)\Psi(x) = 0, \tag{2.1}$$

which is produced by applying the Euler-Lagrange equations to its Lagrangian density, $\mathcal{L}$ [3],

$$\mathcal{L} = \overline{\Psi}(x)(i\gamma^{\mu}\partial_{\mu} - m)\Psi(x). \tag{2.2}$$

This Lagrangian is invariant under global U(1) transformations such as

$$\Psi(x) \rightarrow e^{iQ\theta}\Psi(x), \tag{2.3}$$

where θ is a continuous parameter independent of spacetime position. This global invariance implies the electromagnetic current and charge are conserved. However, the invariance does not hold locally if θ is allowed to depend on spatial position x. Here, the transformation being considered becomes

$$\Psi(x) \rightarrow e^{-iQ\theta(x)}\Psi(x). \tag{2.4}$$

The Lagrangian above is not invariant under this transformation and is therefore not a satisfactory description of nature as it should be independent of spacetime position.

To introduce local gauge invariance, it is necessary to include a gauge vector boson field, $A_\mu(x)$, which transforms as

$$A_\mu \rightarrow A_\mu - \frac{1}{e}\partial_\mu\theta(x). \tag{2.5}$$

A covariant derivative, D_μ can be introduced for convenience, defined as

$$D_\mu\Psi(x) \equiv (\partial_\mu - ieQA_\mu)\,\Psi(x), \tag{2.6}$$

which has transformation properties identical to the field, so

$$D_\mu\Psi(x) \rightarrow e^{-iQ\theta(x)}D_\mu\Psi(x). \tag{2.7}$$

A gauge invariant kinetic term may then be introduced as the field strength tensor, $F_{\mu\nu}$

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x). \tag{2.8}$$

Taking 2.2, adding a vector field $A_\mu(x)$ and a term for $F_{\mu\nu}$, and substituting for D_μ gives the QED Lagrangian density

$$\mathcal{L}_{QED} = \overline{\Psi}(x)(i\gamma^\mu D_\mu - m)\Psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x). \tag{2.9}$$

The vector field introduced here can be identified as the propagator for the electromagnetic force and produces no mass term in the Lagrangian. It is associated with the photon. The addition of a mass term would break the symmetry of the Lagrangian, implying the photon must have zero mass which is consistent with experimental observations.

2.1.3 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory of the strong force, used to describe the interaction of quarks and gluons. Gluons are the gauge bosons that carry the colour charge associated with the force. Colour charge can take values of red, green and blue, as well as three corresponding anti-colours. The gluon can couple to colour-charged particles including itself.

QCD is a non-Abelian[3] gauge theory, based on the $SU(3)_C$ group. The Lagrangian is constructed using a method similar to the QED Lagrangian.

A quark field, ψ, is considered as a triplet of Dirac fields given by

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix}. \tag{2.10}$$

.

The Lagrangian of the quarks coupled to the gluon fields, $\mathcal{L}_{QCD}$, is constructed similarly to the QED Lagrangian as

$$\mathcal{L}_{QCD} = \overline{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a, \tag{2.11}$$

where γ^μ represents the gamma matrices and the covariant derivative, D_μ, is given by

$$D_\mu \equiv \partial_\mu - g_s G^a_\mu T^a, \tag{2.12}$$

where g_s is the strong coupling constant; a is an index running from 1 to 8; and T^a represents the eight generators of the $SU(3)_C$ group. The generators correspond to the eight types of gluon with varying colour charge. This is analogous to the QED photon, where gluons are the mediators of the strong force. The $G^a_\mu(x)$ term refers to the QCD strength tensor, given by

[3] A non-Abelian group is one where the application of a group operator on two or more members of the group is not commutative.

$$G^a_{\mu\nu} \equiv \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^{abc} G^b_\mu G^c_\nu, \tag{2.13}$$

.

where the final term arises from the non-Abelian nature of the $SU(3)_C$ group and represents the coupling of gluon fields to each other. The $SU(3)_C$ structure constants are given by f^{abc}.

The strong coupling constant varies with the energy scale of the interaction, Q^2. At low Q^2, the force becomes very strong leading to *confinement*: as the the distance between two quarks increases so does the force between them, meaning it would take an infinite amount of energy to separate them completely, so quarks are never found in isolation. Conversely, the coupling constant becomes very small in high Q^2 interactions and so the quarks and gluons interact very weakly. This is known as *asymptotic freedom*, and as a result at the very small distances inside hadrons the constituent quarks are essentially free [4]. The hadronisation of quarks is an outcome of this behaviour, observed experimentally as jets of closely-spaced particles in the detector, from which the presence of quarks can be inferred.

2.1.4 Electroweak Theory

The Electroweak (EW) Theory is the unification of the electromagnetic and weak interactions. It is accomplished by requiring that the combined Lagrangian be invariant under local gauge transformations provided by the $SU(2)_L \otimes U(1)_Y$ symmetry group. The generators of the group result in the fundamental quantum numbers weak isospin, I, and weak hypercharge, Y. The weak hypercharge and third component of the weak isospin, I_3, relate to charge, Q, by [5]

$$Q = I_3 + \frac{Y}{2}. \tag{2.14}$$

Up-type quarks and neutrinos have $I_3 = +\frac{1}{2}$ whereas down-type quarks and charged leptons have $I_3 = -\frac{1}{2}$. Quarks cannot decay into other quarks with the same value of I_3 through the weak interaction. So up-type quarks only decay into down-type quarks and vice versa.

Particles can also be categorised by their chirality: the projection of a particle's spin vector in the direction of its momentum vector. If the particle's spin is in the same direction as its momentum its chirality is *right-handed*. If the spin and motion directions are opposite the particle is *left-handed*. Right-handed fermions have $I = 0$ and form singlet states with $I_3 = 0$. Left-handed fermions have $I = \frac{1}{2}$ and form doublets with $I_3 = \pm\frac{1}{2}$. The weak force (W bosons) only interacts with left-handed fermions and right-handed antifermions. In the Standard Model, neutrinos are exclusively left-handed while anti-neutrinos are exclusively right-handed.

The gauge bosons are required to be massless to ensure the local gauge invariance of the Electroweak Lagrangian. As with the QED and QCD Lagrangians, the addition of a mass term to the gauge fields would break the local gauge invariance of the Lagrangian. Through experimental observation, the photon and gluon are indeed massless, but the weak gauge bosons (W^+, W^- and Z^0) are massive. The theoretical framework and experimental observations can be united through the addition of a mechanism to generate mass while preserving the gauge invariance of the Lagrangian. In the Standard Model, this is known as the Higgs Mechanism.

2.1.5 The Higgs Mechanism

The Higgs Mechanism [6] provides the fundamental particles with mass without breaking the gauge invariance of the Lagrangian. This is done by introducing a scalar field, known as the Higgs field, that interacts with the Electroweak boson fields and breaks the symmetry of the $\mathrm{SU}(2) \otimes \mathrm{U}(1)$ group. This field is mediated by the Higgs Boson which was observed experimentally by the ATLAS and CMS experiments in 2012 [7, 8].

The Higgs field is an isospin SU(2) doublet, Φ, of complex scalar fields, Φ^+ and Φ^0

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \begin{pmatrix} \phi^1 + i\phi^2 \\ \phi^3 + i\phi^4 \end{pmatrix}. \tag{2.15}$$

The scalar potential associated with these fields can be described by

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2, \tag{2.16}$$

where μ and λ are constants. When $\mu^2 > 0$, the potential is symmetrical about a minumum at $\Phi = 0$. However, when $\mu^2 < 0$, the potential takes the form shown in Figure 2.1, with an infinite number of minima calculated from

$$|\Phi| = \sqrt{\frac{\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}}, \tag{2.17}$$

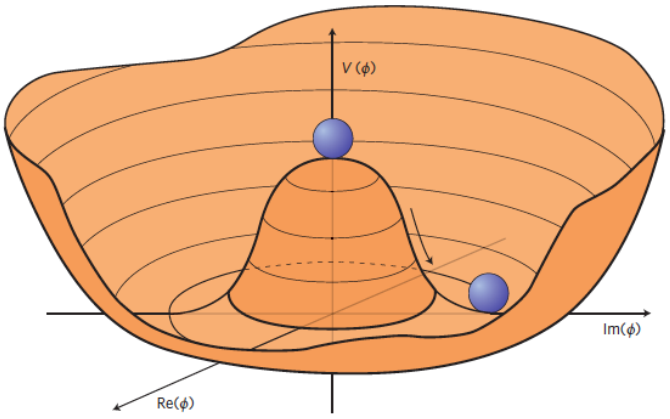

FIGURE 2.1: The shape of the Higgs potential for $\mu^2 < 0$. The indicated motion of the blue ball illustrates Spontaneous Symmetry Breaking which results in a non-zero Vacuum Expectation Value for the Higgs field [9].

where v is the Vacuum Expectation Value (VEV). The solutions correspond to any point on a circle given by

$$|\Phi_0|^2 = \frac{\phi_1 + \phi_2 + \phi_3 + \phi_4}{2} = \frac{v^2}{2}. \tag{2.18}$$

The choice of a particular vacuum is known as the Spontaneous Symmetry Breaking (SSB) of the system, where an initially symmetric system is spontaneously broken by a choice of a specific minimum of the potential. At the time of the Big Bang, the universe was in a state where the VEV was 0, but this was unstable.

The vacuum is chosen so that $\phi_1 = \phi_2 = \phi_4 = 0$. Using the Goldstone Theorem, each of the four fields has an associated Goldstone boson, three of which are massless and unphysical and the fourth is the Higgs Boson. A gauge can be chosen in which the three unphysical fields (ϕ_1, ϕ_2, and ϕ_4) are eliminated. This is known as the Unitary Gauge and tranforms Equation 2.15 into

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \tag{2.19}$$

where H is the neutral scalar Higgs field. This also gives rise to mass terms for the W^+, W^- and Z^0 bosons.

The Electroweak interaction propagates through four gauge fields: B_μ, W_μ^1, W_μ^2 and W_μ^3. B_μ is required for invariance to be maintained under electromagnetic transformations, while the W_μ fields are required to maintain invariance under weak transformations. These gauge fields are related to the physically observable vector bosons by

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp W_\mu^2 \right), \tag{2.20}$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu, \tag{2.21}$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu, \tag{2.22}$$

where θ_W is the weak mixing angle, also known as the Weinberg Angle. This is related to the weak isospin coupling constant, g, and the weak hypercharge coupling constant, g' by

$$\tan\theta_W = \frac{g'}{g}. \tag{2.23}$$

The Lagrangian for the four Electroweak fields can be written to include the interaction with the Higgs field

$$\mathcal{L}_{Higgs} = \left(D_\mu \Phi\right)^\dagger \left(D^\mu \Phi\right) - \mu^2 \Phi^\dagger \Phi - \lambda \left(\Phi^\dagger \Phi\right)^2, \tag{2.24}$$

which can be written in the Unitary Gauge as

$$\begin{aligned}\mathcal{L}_{\text{Higgs}} = &\frac{1}{2}\partial_\mu H \partial^\mu H + \frac{1}{4}g^2\left(H^2 + 2vH + v^2\right)W_\mu^+ W^{-\mu} \\ &+ \frac{1}{8}\left(g^2 + g'^2\right)\left(H^2 + 2vH + v^2\right)Z_\mu Z^\mu \\ &- \mu^2 H^2 - \frac{\lambda}{4}\left(H^4 + 4vH^3\right).\end{aligned} \tag{2.25}$$

The mass terms corresponding to the $W^\pm$ and Z^0 bosons are given by:

$$M_W = \sqrt{\frac{1}{4}g^2 v^2} = \frac{1}{2}gv,$$

$$\begin{aligned}M_Z &= \sqrt{\frac{1}{4}\left(g^2 + g'^2\right)v^2} = \frac{1}{2}\left(g^2 + g'^2\right)^{1/2} v \\ &= \frac{1}{2}\frac{gv}{\cos\theta_w} \\ &= \frac{M_W}{\cos\theta_w},\end{aligned} \tag{2.26}$$

where the relations $(\cos^2\theta_w + \sin^2\theta_w) = 1$ and $g' = g\tan\theta_w$ are used to express the Z^0 mass in terms of the $W^\pm$ mass and the weak mixing angle. The charged fermions aquire mass through the addition of a further term, $\mathcal{L}_{yukawa}$, used to describe so-called Yukawa coupling.

2.1.6 Yukawa Coupling

A coupling between the Higgs field and fermions can be added to the Standard Model Lagrangian to give the fermions mass. A contribution, $\mathcal{L}_{yukawa}$, can be used to describe the coupling between the fermions and the Higgs doublet:

$$\mathcal{L}_{\text{yukawa}} = -g_f \overline{\Psi}_f \Psi_f \Phi. \tag{2.27}$$

An expansion around the vacuum expectation value of the Higgs doublet, ν, gives

$$\mathcal{L}_{yukawa}(f) = -\frac{1}{\sqrt{2}}g_f \nu \overline{\Psi}_f \Psi_f - \frac{1}{\sqrt{2}}g_f \overline{\Psi}_f \Psi_f H. \tag{2.28}$$

Here, the first term represents a mass term where

$$m_f = \frac{g_f \nu}{\sqrt{2}}. \tag{2.29}$$

The second term represents the coupling of the fermion to the Higgs field, given by

$$\frac{2m_f}{\nu}. \tag{2.30}$$

The Yukawa coupling constant of the fermion is given by g_f. These coupling constants are not predicted by the Standard Model and so can be set to match the fermion masses observed in experiment. The coupling with the Higgs field is proportional to the fermion's mass.

2.1.7 Quark Mixing and CP Violation

Weak and strong mass eigenstates of quarks are not the same and so quark flavour and generation are not conserved in weak decays. The relationship between the flavour eigenstates (d', s' and b') and the mass eigenstates (d, s, and b) is shown by:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.31}$$

where V_{CKM} is a 3×3 unitary matrix used to describe the mixing between quark generations. It is known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [10], given by:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{2.32}$$

Each element describes the mixing between two quark flavours: for example, V_{tb} is the coupling between a t quark and a b quark. The matrix can be written in terms of three quark mixing angles (θ_{12}, θ_{13} and θ_{23}) and a single CP-violating phase, δ:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.33}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. If all three angles were equal to zero, the matrix would become the identity matrix and quark mixing would be prevented from occurring. Results from experimental fits for the magnitudes of all nine CKM elements are shown below [1].

$$V_{\text{CKM}} = \begin{pmatrix} 0.97446 \pm 0.00010 & 0.22452 \pm 0.00044 & 0.00365 \pm 0.00012 \\ 0.22438 \pm 0.00044 & 0.97359^{+0.00010}_{-0.00011} & 0.04214 \pm 0.00076 \\ 0.00896^{+0.00024}_{-0.00023} & 0.04133 \pm 0.00074 & 0.999105 \pm 0.000032 \end{pmatrix} \tag{2.34}$$

Current understanding of the matrix can also be shown in graphical form as the Unitarity Triangle. The sides and angles of the triangle represent various ratios of the CKM parameters. Measurements from a number of interactions at multiple experiments contribute to fits used to constrain the triangle and so the CKM parameters. An example of a CKM fit is shown in Figure 2.2 where different contributions to the fit are represented by the various coloured regions. The contributions made by LHCb are outlined in Reference [11], along with predictions of how future CP violation measurements made at LHCb are expected to further constrain the model.

The phenomenon of CP-violation is not fully understood but must be parameterised in the Standard Model in order for precise theoretical predictions to be made.

2.1.8 Summary

The Standard Model is the current theory for describing the fundamental particles of nature and their behaviour. However, it does not fully explain all of the physical observables of matter such as the mechanism responsible for neutrino mass; the large range in mass of the fundamental particles; or the composition of dark matter or dark energy. It

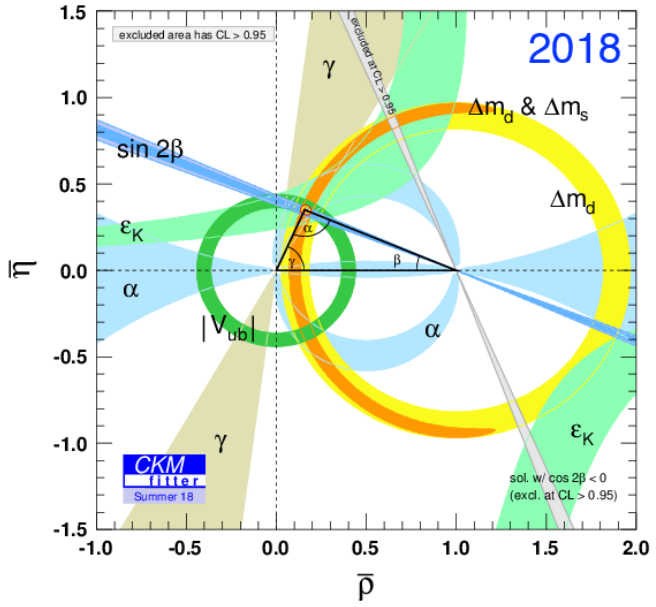

FIGURE 2.2: A CKMfitter Collaboration fit to the Unitarity Triangle with individual constraints from various measurements superimposed [11].

is important to make experimental tests of the Standard Model to confirm our understanding of particle physics, and to lead us toward searches for new physics beyond the SM which has yet to be discovered.

2.2 Testing the Standard Model at Hadron Colliders

Collisions between high-energy hadrons at particle accelerators can be used to test and validate predictions made by the Standard Model. At the Large Hadron Collider (LHC), beams of protons are accelerated to close to the speed of light in order to produce high-energy collisions. These collisions produce particles that are not naturally observable in nature, which quickly decay into various *final state* particles. Deposits left by these final state particles in particle detectors can be used to reconstruct collision events. Such observations can then be used to make experimental comparisons with theoretical SM predictions, provide clues for as yet undiscovered physics beyond the SM, or confirm the existence of previously unseen particles, such as the Higgs boson.

2.2.1 Hard Scattering and the Underlying Event

The composite nature of the particles colliding at hadron colliders leads to complex event structures. Two distinct processes of interest can be defined: the hard scatter and the underlying event.

Hard scattering refers to the initial high energy interactions between partons during collisions of hadrons. These are typically characterised by a large momentum transfer, high transverse momentum or large mass scale. This is the process responsible for the production of new particles, such as top quark pairs. Interactions such as this can be calculated to high precision using perturbative techniques.

The underlying event is any activity observed in a hadron collider event which is not a product of the primary hard scattering process. The composite nature of the colliding particles means that the partons not involved in the primary process can still interact. These interactions are softer than the primary interaction, but contribute a non-negligible contribution to the overall event and its multiplicity. Such multiple interactions can be evaluated through ordering the scattering in a declining p_{T} scale and so arranging the interactions from hardest to softest scattering. Correlations can also be taken into account, such as those arising from flavour or momentum conservation. The underlying event can include both initial and final state radiation, beam-beam remnants and multiple parton interactions. Events where multiple collisions occur per beam crossing (known as pile-up) also contribute to the event structure. Understanding the underlying event is important for event simulation and in selecting the desired signals from data. The underlying event is modelled using Pythia8 simulation tuned to data.

Calculations used to model hadronic interactions must include both the hard scattering process and any soft processes involved in the underlying event. The relationship between the overall cross-section (or hadronic cross-section) and the cross-section of the hard process (or partonic cross-section) can be understood by modelling the internal structure of the colliding hadrons. Experimental data can be used to calculate Parton Distribution Functions (PDFs) to model this internal structure and so predict the likelihood of specific interactions occuring between colliding hadrons.

2.2.2 Parton Distribution Functions

Parton Distribution Functions are the probability densities for finding partons with momentum fraction, x, at a given energy scale, Q^2, inside hadrons. They are process-independent, non-perturbative functions. PDFs must be extracted from experimental data at a given scale, due to limitations in current lattice QCD calculations. They can then be transformed to different energy scales using the DGLAP equations [12–14] and used in theoretical calculations.

PDFs can be extracted from measurements performed at the LHC using the Factorisation Theorem. This breaks down the hadronic cross-section into the product of a partonic-level cross-section and a parameterisation of the colliding hadrons, which is determined by the PDFs. The production of an object, X, can be determined from:

$$\sigma_{AB\to X} = \int dx_a dx_b f_{a/A}\left(x_a, Q^2\right) f_{b/B}\left(x_b, Q^2\right) \hat{\sigma}_{ab\to X}, \tag{2.35}$$

where A and B are the colliding hadrons and a and b are the partons involved in the hard scattering process. The functions $f_{a/A}\left(x_a, Q^2\right)$ and $f_{b/B}\left(x_b, Q^2\right)$ are the PDFs for the partons as a function of the momentum fraction they carry ($x_{a/b}$) and the energy transferred in the interaction (Q^2). Finally, $\hat{\sigma}_{ab\to X}$ is the partonic level cross-section, which represents the cross-section for the production of X from partons a and b. It is calculated peturbatively.

Any prediction of physics observable at the LHC requires knowledge of PDFs as the modelling of expected processes relies on precise knowledge of the proton structure. In turn, specific measurements made at the LHC can help constrain PDFs and resolve disagreements between different PDF modelling groups. These measurements include Drell-Yan processes, W/Z boson cross-sections, diboson production and top quark production [15].

The PDF sets used in the following analyses are produced by the CTEQ [16] and NNPDF [17] groups. CTEQ uses global analysis to determine PDFs by fitting theoretical QCD cross-sections to existing hard scattering data. NNPDF determines PDFs using Neural Networks to construct Monte Carlo representations of the PDFs and their uncertainties.

2.3 The Top Quark

The primary analysis presented in this thesis is a test of the Standard Model using top quark production. This is performed by measuring how often pairs of top quarks are produced at the LHCb experiment and comparing that measurement with the theoretical prediction. This Section begins with some top physics context and the motivation for studying the top quark, followed by a summary of top quark production and decay modes.

2.3.1 Motivation

The existence of the top quark was first theorised after the discovery of CP violation in 1964 [18]. Until then, experimental particle physics results were adequately explained by a two generation Standard Model, but the subsequent development of the CKM matrix introduced quark mixing parameters that required a third generation of particles. This led to searches for, and the subsequent discovery of, two quarks (top and bottom) as well as the tau lepton and its corresponding neutrino.

A prediction of the top quark mass was made using electroweak constraints in 1994, anticipating a particle with a mass of $\sim$175 GeV [19]. The experimental discovery of the top quark then followed in 1995 at Fermilab's CDF and D0 experiments [20, 21].

The top quark's mass is currently measured to be 172.9$\pm$0.4 GeV [1]. This is well above the scale of the other quarks and makes the top quark the heaviest of all elementary particles discovered to date. Because of this high mass, the top quark is thought [22] to couple strongly to new physics in a number of models. It is therefore a natural probe of physics beyond the standard model, for example at the Minimal Supersymmetric Standard Model scale [22].

The top quark is also the only quark with a Yukawa coupling to the Higgs boson of close to unity. A better understanding of top quark physics could potentially help to explain why the mass of the Higgs boson is considerably lower than predicted [22]. It has also been postulated that the top quark's Yukawa coupling could be used to determine the energy scale at which new physics will be found at the hadron colliders of the future, due to its contribution to the scale of the Higgs self-coupling [23].

Top quark production was first observed at the LHC by the ATLAS and CMS experiments during Run-I [24]. Combined measurements of the top quark pair-production cross-section by these experiments at 7 and 8 TeV produce a degree of uncertainty low enough to challenge theoretical predictions [22], making Run-II an exciting time for the study of top physics. Further measurements of the cross-section have been made in Run-II and are consistent with theoretical predictions. The current status of ATLAS and CMS measurements is summarised in Figure 2.3, with the most precise measurement made in the μe channel.

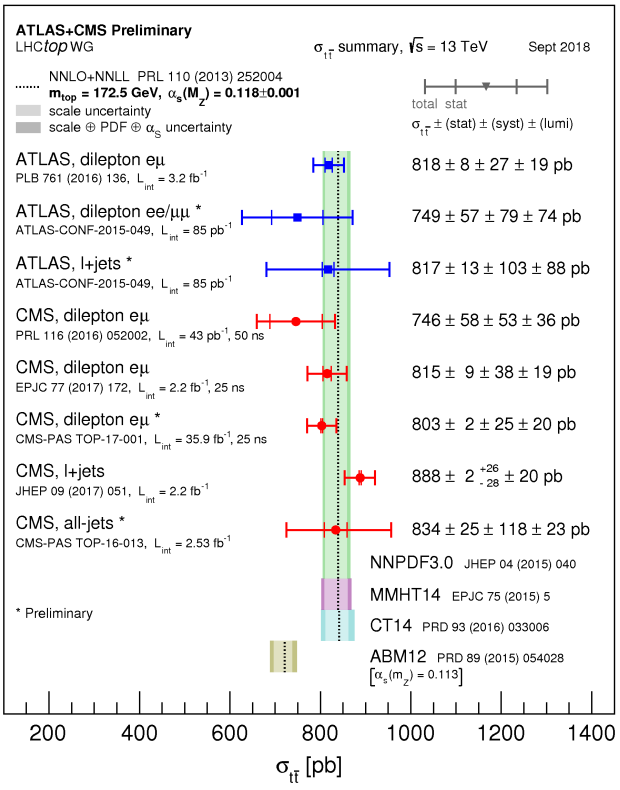

FIGURE 2.3: A summary of $t\bar{t}$ production cross-section measurements at the ATLAS and CMS experiments [25].

Also during Run-I, the LHCb experiment became the fifth experiment to observe top production and the first to do so in the forward region [26]. LHCb's unique forward acceptance allows for the probing of a phase space that is inaccessible to the other experiments currently measuring top quark production. Measurements at LHCb are able access a higher value of Bjorken-x than the central detectors. There are large uncertainties in the proton PDFs in this region and so precise measurements of quark production at LHCb can be used to constrain them.

2.3.2 Production of Top Quark Pairs

At the LHC, top quarks are produced through quark-antiquark annihilation and gluon-gluon fusion with the latter mechanism accounting for around 90% of all LHC top quark production in Run-II[4][1]. Measuring the production rates of both single top and top pair production can therefore contribute substantially to the modelling of the gluon PDF. The production ratio of the top quark is shown in Figure 2.4, with LHCb's unique forward acceptance highlighted in grey.

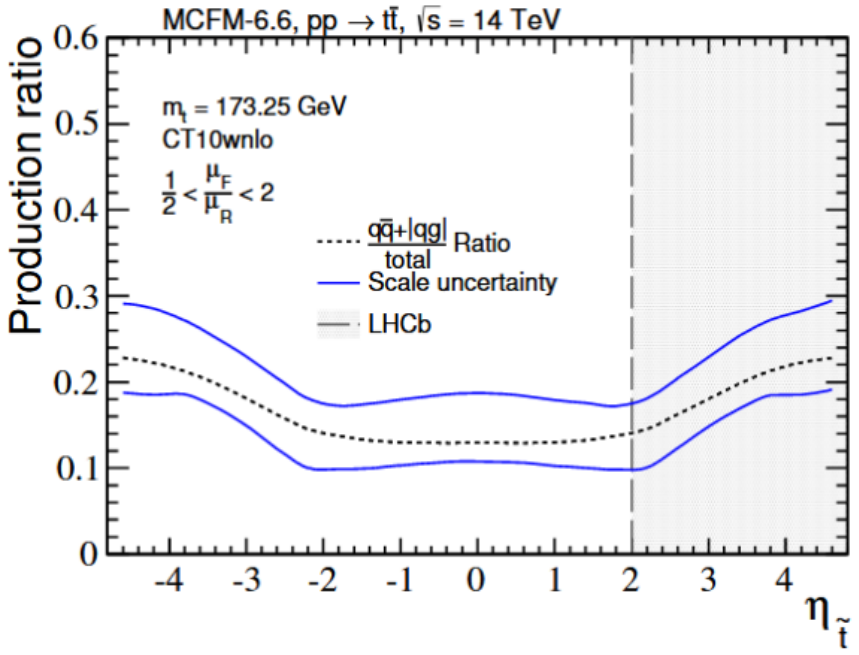

FIGURE 2.4: The production ratio of top quarks as a function of top pseudorapidity at $\sqrt{s}$ 14 TeV [27].

The production meachanisms for top quark pairs can be represented using Feynman diagrams. Feynman diagrams present a graphical representation of the mathematical expressions that describe the behaviour of particles. Lines are used to represent particle currents with time running from left to right. The points where lines meet are called vertices and represent interactions.

Leading order (LO) Feyman diagrams display the simplest possible interaction to produce a given final state. Examples of LO diagrams for $t\bar{t}$ production are shown in Figure 2.5. Feynman diagrams can be read from left to right for initial to final states. For example,

[4]The production ratio is dependent on the centre of mass energy.

in the LO example diagram (a) in Figure 2.5 two gluons from colliding protons fuse together to create a higher energy gluon which then decays to a $t\bar{t}$ pair.

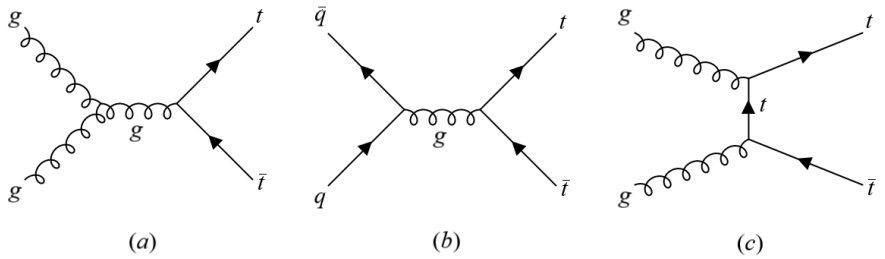

FIGURE 2.5: Feynman diagrams for the production of top quark pairs in leading order processes at the LHC: (a) gluon fusion, (b) pair creation from $q\bar{q}$ annihilation, (c) t-channel gluon fusion.

Next-to-leading order (NLO) diagrams are used to represent more complex mechanisms with extra interactions. Examples of NLO diagrams for $t\bar{t}$ production are shown in Figure 2.6. Here, diagram (a) shows a $q\bar{q}$ annihilation where one of the incident partons undergoes initial state radiation.

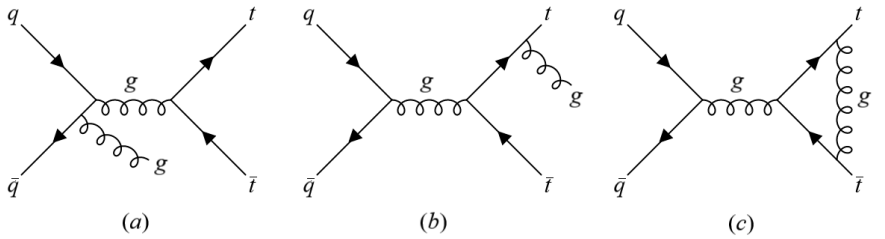

FIGURE 2.6: Examples of next-to-leading order production meachanisms for top quark pair production.

Each vertex in a Feynman diagram causes a process to be rarer by an order of the strong coupling constant, α_s. In the examples provided, diagram (a) in Figure 2.5 has a cross-section of order α_s^2, while diagram (a) in Figure 2.6 has a cross-section of order α_s^3 and diagram (c) in the same figure has a cross-section of α_s^4. Therefore, while there are significantly more possibilities for drawing NLO diagrams than LO, the NLO contribution to the overall production cross-section is generally reduced.

2.3.3 Top Quark Decay

The top quark's high mass makes it very unstable. It therefore has a short lifetime of $\sim 0.5 \times 10^{-24}$s which is much shorter than that of any other quark [1]. This gives it the

unique property of decaying before any hadronisation can occur, which means that top provides a rare opportunity to study a bare quark directly.

Top quarks decay almost exclusively to a bottom quark and W boson [1] as determined by the V_{tb} term of the CKM matrix. As b-quarks hadronise, the presence of a b-quark in a hadron collider top quark event must be inferred from the presence of a b-jet. The W boson can be identified through its subsequent decay to $l\nu$ states (33%) or $q\bar{q}$ pairs (67%) where the presence of the quarks is again inferred through the detection of jets.

Identifiable final states for top pair production can therefore range from two b-jets and two charged leptons to two b-jets and four jets. Here, the former is known as the *dilepton* channel and the latter as the *all-hadronic* channel.

Specific consideration must also be given to cases where the W boson decays to tau leptons. This is because the tau decays to either an electron or muon only ∼35% of the time with the remaining ∼65% of tau decays being hadronic. For this reason, the tau lepton decay modes of the W boson are sometimes considered as hadronic decay channels.

A graphical summary of the likelihood of the top quark pair decay channels can be seen in Figure 2.7.

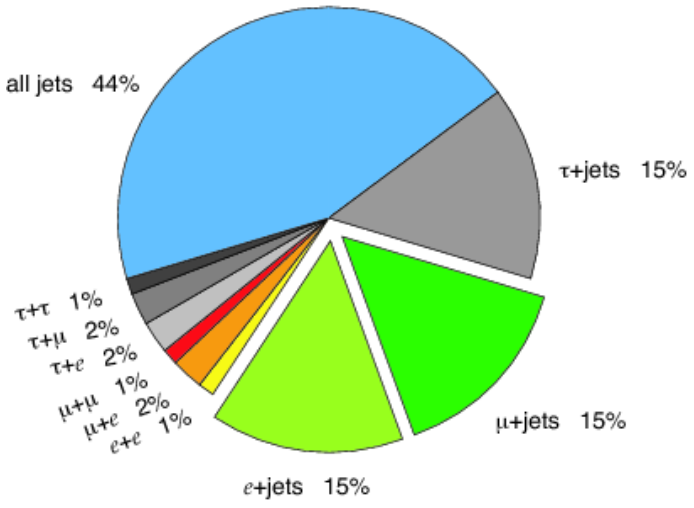

FIGURE 2.7: The different branching ratios of a top quark pair, referring to the decay modes of the two W bosons arising from the top quark decays [28].

In the analysis presented in Chapter 5, the top quark pair production cross-section is measured in the μeb chanel.

Chapter 3

Experimental Environment

In this chapter, the Large Hadron Collider and LHCb experiment are described. First, the journey of the colliding protons through the accelerator complex is outlined, followed by a discussion of the LHC running performance. After this, there is a description of the LHCb experiment, along with a more detailed look at each of the the sub-detectors used to make measurements.

3.1 The Large Hadron Collider

The Large Hadron Collider is a proton-proton collider at the European Organisation for Nuclear Research (CERN), near Geneva, Switzerland. It is housed in a circular tunnel with a 27 km circumference approximately 100 m beneath the Franco-Swiss border. Two beams of protons travel around the ring in counter-rotating directions. There are four main experiments situated at various collisions points around the beam, used to collect data from the proton interactions: ATLAS, CMS, ALICE and LHCb. A diagram of the ring is shown in Figure 3.1.

3.1.1 The Accelerator Complex

The protons involved in collisions at the LHC are first accelerated by a series of smaller accelerators. They begin their journey as hydrogen gas, which is passed through an electric field to strip off the electrons. They then enter the accelerator chain, starting

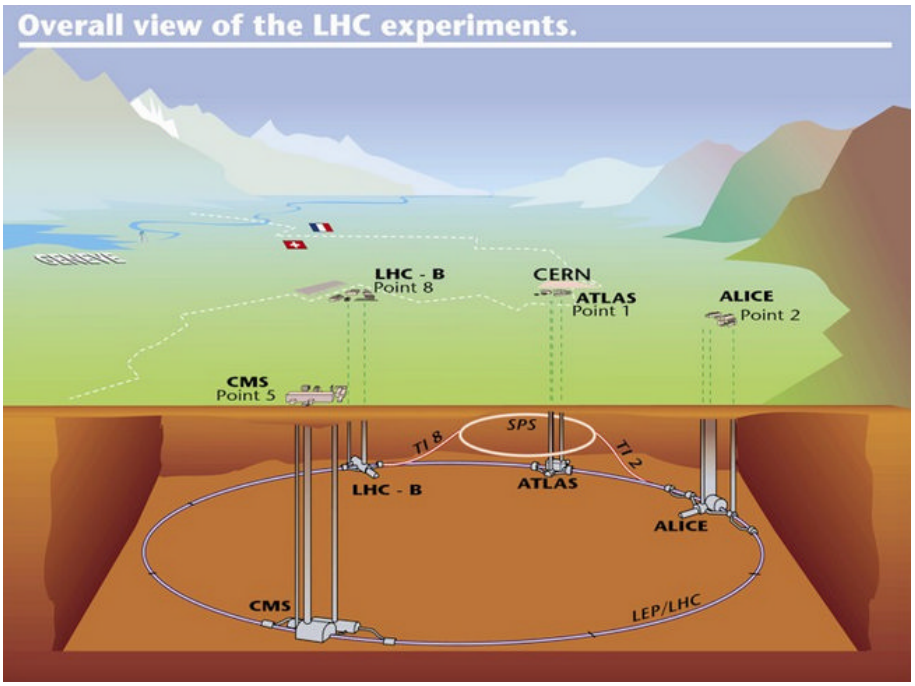

FIGURE 3.1: The LHC ring with the CERN site and four experiments labelled [29].

with the Linear Particle Accelerator (LINAC) which accelerates the protons to an energy of 50 MeV. The protons are then injected into the Proton Synchrotron (PS), which boosts them to 25 GeV. Next, the Super Proton Synchrotron (SPS) brings the protons up to the injection energy of 450 GeV and they are transferred to the LHC where they are accelerated to their collision energy. The accelerator complex is shown in Figure 3.2.

The LHC accelerates protons using radio-frequency (RF) cavities [31]. These are metallic chambers containing oscillating electromagnetic fields. The energy a proton gains in an RF cavity depends on its time of arrival. This results in the proton beam being split into discrete groupings or *bunches*. Each beam is designed to hold 2808 bunches of 10^{11} protons. The bunches are separated by approximately 25 ns, resulting in a collision rate of 40 MHz.

Once the protons have reached the desired energy, magnets are used to direct and focus them onto the point of collision at each experiment. The detectors are able to record collision data from each beam for around 12 hours until the beam is *dumped*. This involves redirecting the beam into a graphite block designed to absorb its energy and occurs when the the number of protons left in the beam gets too low or if there are technical issues with the beam.

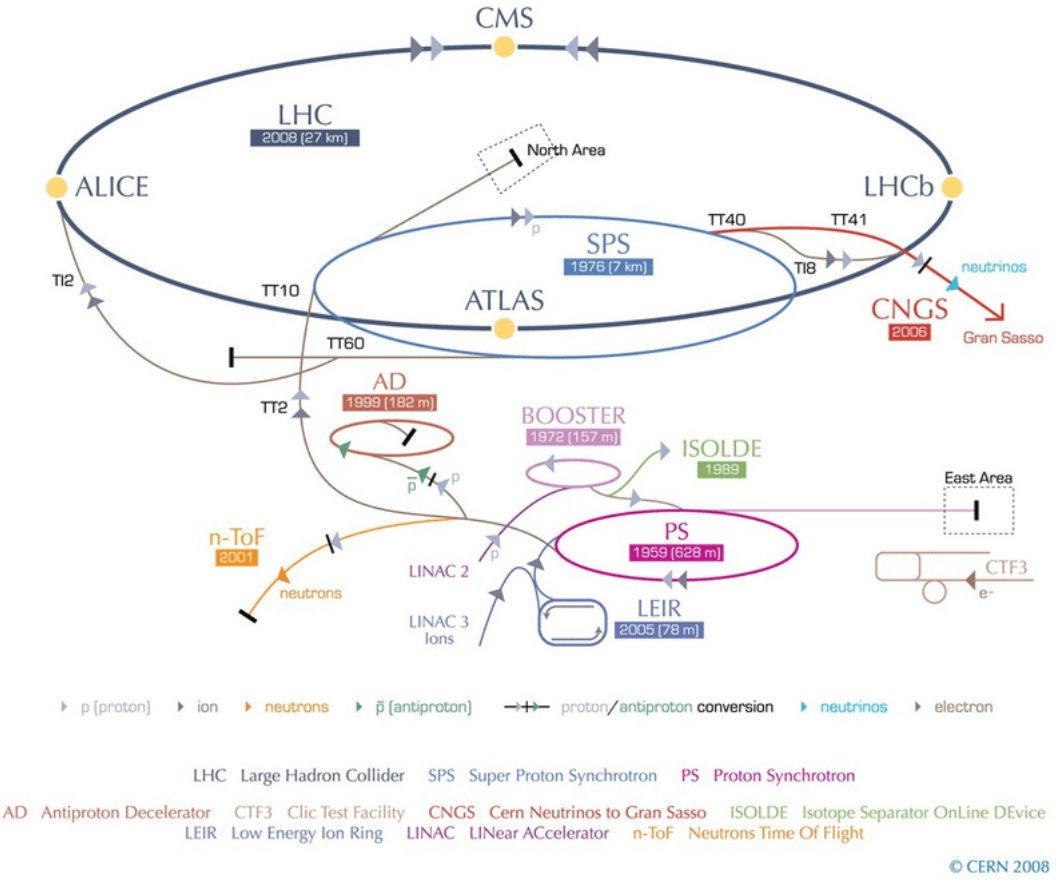

FIGURE 3.2: The LHC accelerator complex [30].

The time between the protons entering the accelerator complex and the beam being dumped is known as a *fill*. Each experiment segments a given fill into a number of *runs* depending on its own criteria. A new run can be started for various reasons including changes to configuration settings or a given sub-detector being enabled/disabled.

3.1.2 Running Performance

The reach of the physics analysis undertaken at the LHC is determined by the centre of mass energy of the collisions and the luminosity. The collision energy determines what particles can be produced and the luminosity dictates how often such events occur.

The data presented in this thesis was collected during the LHC's second run from 2015 to 2018 when the centre of mass energy of the collisions was 13 TeV. This is lower than the design energy of 14 TeV, due to an incident that occurred during Run-I which caused damage to the machine [32] and the subsequent cautious running of the LHC.

The instantaneous luminosity, $\mathcal{L}$, is a measure of the number of collisions that take place within a detector per cm^2 per second. It is determined by:

$$\mathcal{L} = f\frac{N_1 N_2}{4\pi\sigma_x\sigma_y} \tag{3.1}$$

where f is the bunch crossing frequency, N_1 and N_2 are the number of protons in the bunches in each beam, and σ_x and σ_y describe the dimensions of the overlapping beam profile. Typical values for these components can be found in Reference [33].

The LHC was designed to operate at an instantaneous luminosity of 10^{34} $\text{cm}^{-2}\text{s}^{-1}$ for ATLAS and CMS, and a lower instantaneous luminosity of 10^{32} $\text{cm}^{-2}\text{s}^{-1}$ for LHCb. Because of this, LHCb sees $\mathcal{O}(1)$ pp collisions per bunch crossing, which is significantly lower than ATLAS and CMS. This reduction is crucial to LHCb's b-physics program as it removes ambiguity in determining the vertex from which a given B meson originated. It also brings sub-detector occupanies down to an acceptable level and helps reduce radiation damage within the detector.

LHCb's lower instantaneous luminosity is enabled by reducing the overlap region of the colliding proton beams. This allows LHCb to perform a technique known as luminosity levelling [34], where the overlap region is adjusted throughout the data taking period to keep the instantaneous luminosity constant. A comparison of the instantaneous luminosity delivered to the ATLAS, CMS and LHCb experiments in a typical fill can be seen in Figure 3.3.

The integral of the instantaneous luminosity over time is known as the integrated luminosity. This is a measurement of the amount of data collected. The integrated luminosity recorded by LHCb during its lifetime so far can be seen in Figure 3.4. In Run-II, a total integrated luminosity of 5.9 fb^{-1} was delivered to LHCb, as shown in Figure 3.4.

There are two methods used to determine luminosity at LHCb [33]. Van der Meer scans are used to determine the beam profile by measuring the collision rate as a function of beam displacement as one of the beams is swept sideways across the other [37]. A beam gas method is also used, in which small amounts of gas are allowed into the beam vacuum [38]. The beam profile is measured by reconstructing vertices arising from beam-gas interactions. The combined measurement of the beam profile can then be used to determine the luminosity from Equation 3.1. For 2015 and 2016 data the uncertainty

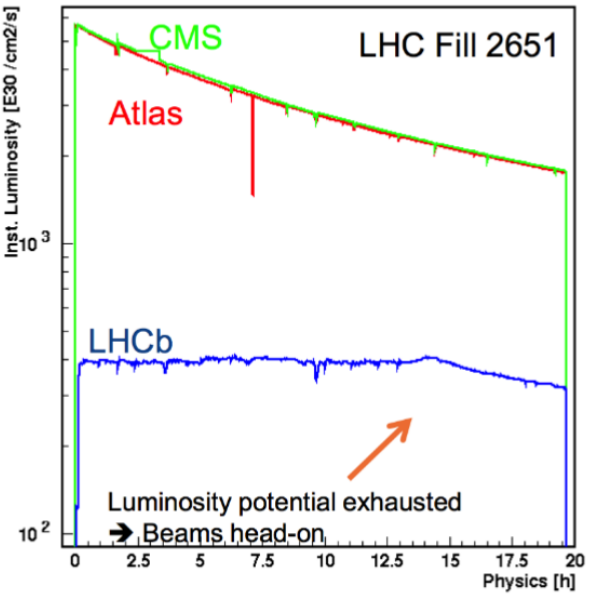

FIGURE 3.3: The instantaneous luminosity over time at ATLAS, CMS and LHCb during fill 2651 of the LHC in May 2012 [35].

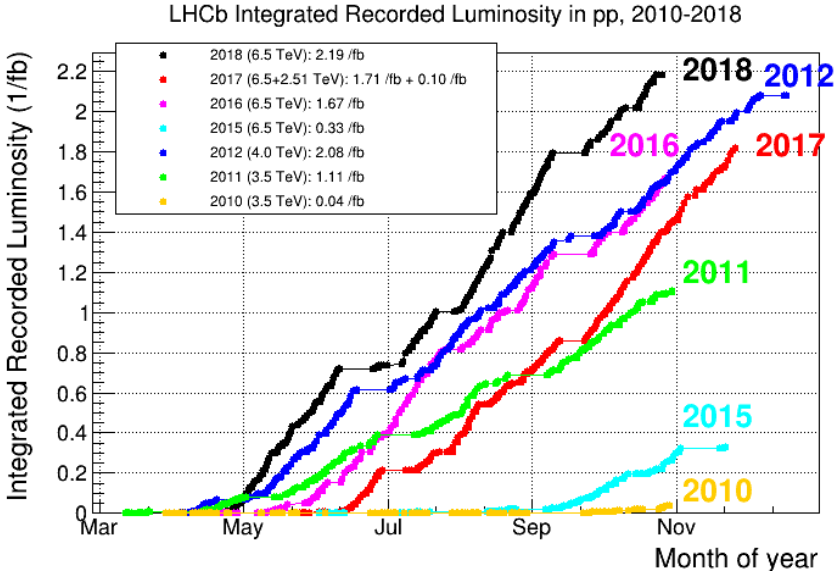

FIGURE 3.4: The integrated luminosity recorded by LHCb per year during Run-I and Run-II [36].

assigned to the luminosity is 3.9%[39]. At the time of writing the luminosity calibration for 2017 and 2018 data is not complete and so a larger uncertainty is assigned for these years for the purpose of the following analyses, as discussed in Section 5.2.

3.2 The LHCb Experiment

The LHCb detector is a single-arm spectrometer in the forward region. The experiment was originally designed to make precision measurements of decays of particles containing b quarks, which are predominantly produced in the forward or backward directions in pp collisions, as shown in Figure 3.5.

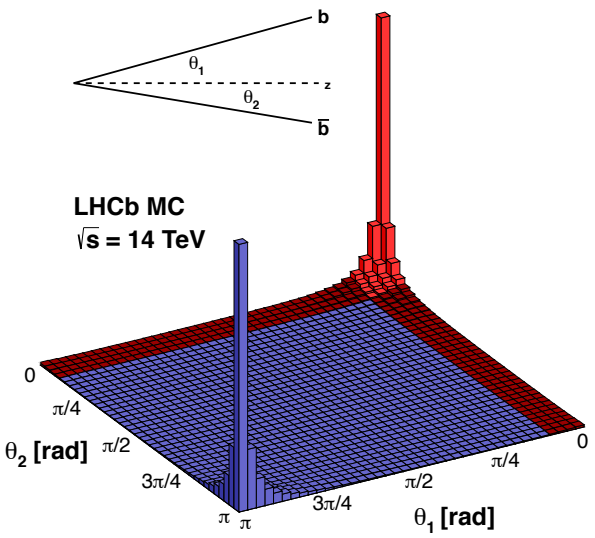

FIGURE 3.5: The angular direction of $b\bar{b}$ production in simulated collisions at 14 TeV, with the LHCb acceptance indicated in red [40].

A diagram of the LHCb experiment and its various sub-detectors can be seen in Figure 3.6. The proton beams collide at the Interaction Point within the Vertex Locator (VELO) which makes the first measurements of charged tracks used in reconstruction and vertexing. Particles produced in the interaction then travel through the first of two particle-identifying RICH detectors and the silicon TT tracker. A dipole magnet is used to bend charged particle tracks which are then observed in three further tracking stations, T1-T3. Further particle identification is then performed using the second RICH detector, the calorimeter system and five muon chambers, M1-M5. The methods used to

reconstruct an event from deposits left throughout the detector by final state particles are discussed in Section 4.

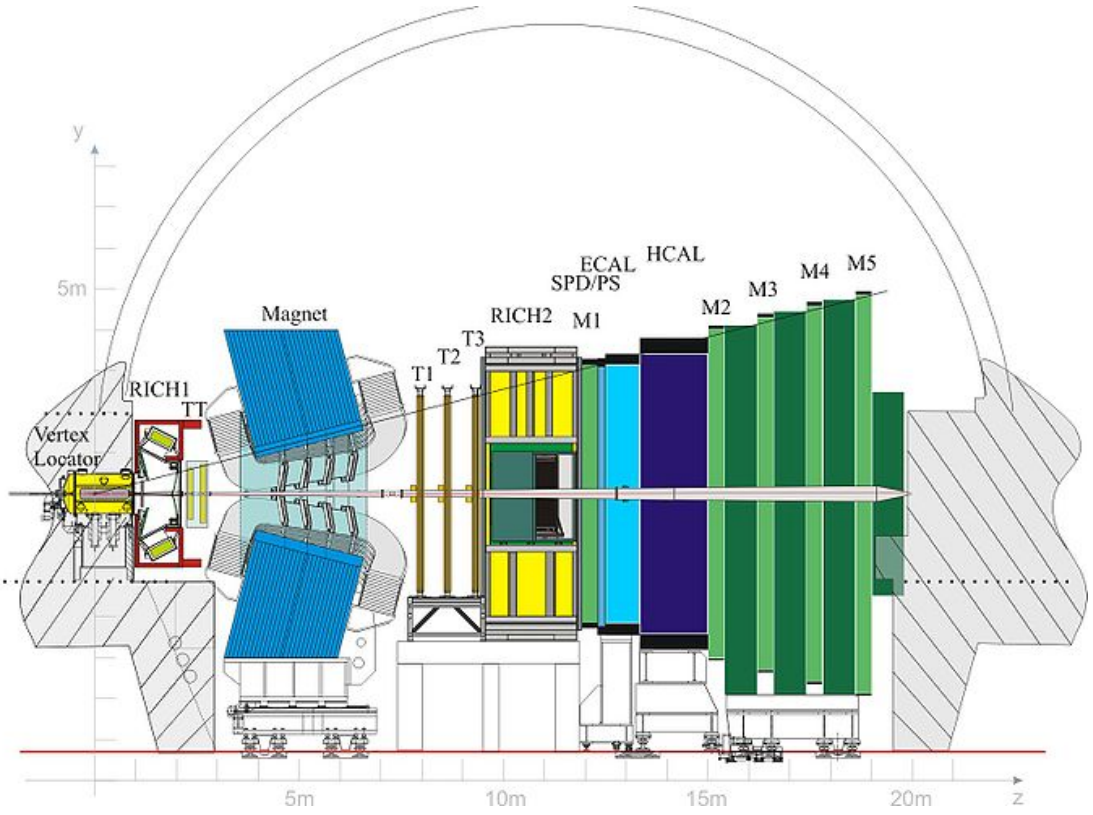

FIGURE 3.6: The LHCb experiment with each detector sub-system labelled. Particles travel from left to right through the detector [41].

Particles travel forwards through the detector from the Interaction Point within the VELO towards the muon stations in the direction of the $+z$ axis. This is from left to right in Figure 3.2. From the same perspective, the magnetic field exerted by the dipole magnet points either *up* or *down* in the y direction and bends charged particle tracks into of out and the page in x. For this reason, the $x-z$ plane is also referred to as the *bending plane* and the $y-z$ plane as the *non-bending plane*. The $x-y$ plane is known as the transverse plane.

This geometry can also be defined using $\eta-\phi$ space, with ϕ being the azimuthal angle around the z axis ranging from $-\pi$ to π radians, and the pseudorapidity, η defined as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right), \tag{3.2}$$

where θ is the polar angle with respect to the beampipe which runs along the z axis. LHCb is fully instrumented in a forward angle between 10 - 300 mrad in the bending plane and 10 - 250 mrad in the non-bending plane. This is equivalent to an acceptance of $2 < \eta < 4.5$.

The positions of sub-detectors and particle tracks are often referred to as being *up-* or *downstream*. This refers to the direction of the flow of particles through the detector with the most upstream system being the VELO and the most downstream the muon stations.

Each of the sub-detectors that make up the LHCb experiment is described in further detail below.

3.3 Tracking System

The Tracking System is made up of the Vertex Locator (VELO) and four tracking stations, one positioned before (TT) and three after (T1-T3) the dipole magnet. The VELO, upstream tracker and inner sections of the downstream trackers utilise silicon detector technology. The outer regions of the downstream trackers use drift-tubes to detect particle tracks. The layout of the tracking stations can be seen in Figure 3.7.

3.3.1 Vertex Locator

The LHCb Vertex Locator (VELO) is a series of silicon sensor modules used to precisely measure particle trajectories close to the interaction point. These sensors are a form of semiconductor dectectors in which charged particles travelling through the silicon create ionisation currents by freeing electrons and holes. An electric field applied across the sensors causes these free charges to travel to electrodes where they create an electrical pulse that can then be amplified and detected. Signals produced throughout the detector can be combined to map a charged particle's trajectory. These measurements are then used to locate and separate primary and secondary vertices.

The VELO is made up of two rows of half-moon modules, covering $\sim$1 m in z around the interaction point [42]. During injection, the VELO is open and the modules are retracted to a safe distance of 29 mm. The VELO is then closed around the beam when

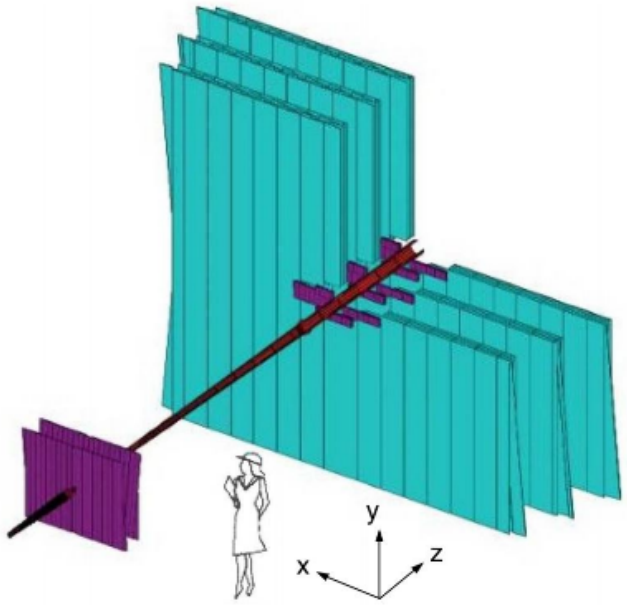

FIGURE 3.7: Layout of LHCb tracking stations. The TT is on the left of the image, T1-T3 are on the right. Silicon trackers are displayed in purple and drift-time trackers in blue. The VELO is situated downstream of the TT station at a lower position on the z axis [41].

it is stable and the active detector areas are ∼8 mm from the beam during data-taking. A diagram of the VELO can be seen in Figure 3.8.

Each of the VELO's 42 modules is comprised of two back-to-back silicon sensors with R and ϕ geometry respectively as shown in Figure 3.9. Three dimensional co-ordinates can be determined by combining a module's (r, z) and (ϕ, z) measurements.

As the VELO is opened and closed for every new beam, its position must be well known to ensure accurate measurements of particle trajectories. A software alignment is performed after each closing before data taking begins [43]. Initial positions for each module are taken from a survey after the initial construction with a precision of better than 10 μm. The alignment then consists of three stages: aligning the r and ϕ sensors of each module with respect to each other; determining the positions of the modules in each VELO half relative to each other; and aligning the two halves. These methods are performed by minimising track residuals, defined as the shortest distance from the reconstructed track trajectory to the hits left by the track in the detector.

The VELO is used to locate the points from which particles created in particle collisions originate, known as primary vertices (PVs). It is also useful in helping to determine the

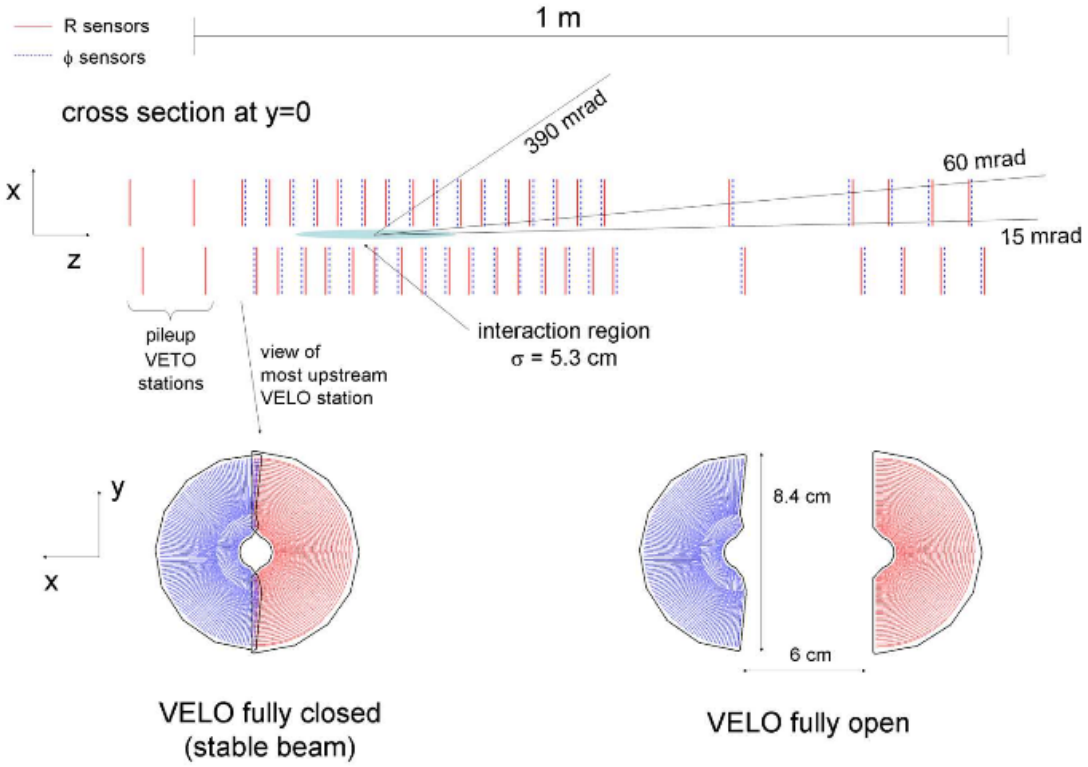

FIGURE 3.8: The layout of the LHCb VELO showing the position of the 42 modules around the interaction region. Fully closed and fully open module positions are also shown [41].

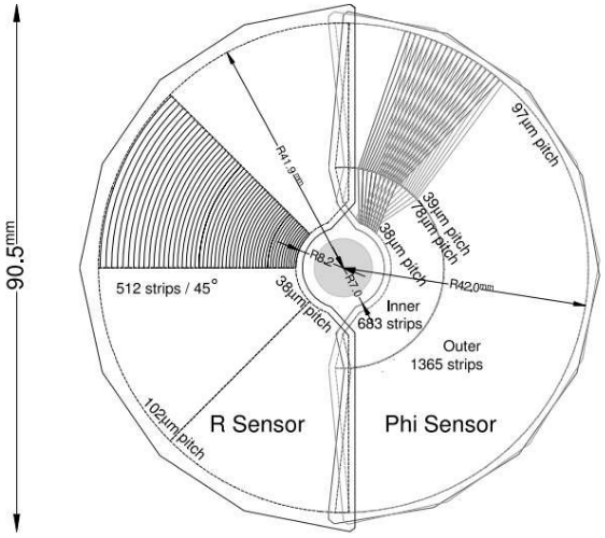

FIGURE 3.9: Detailed view of a pair of closed VELO modules with R and ϕ geometry [41].

lifetime of particles by measuring their impact parameter, the distance approach of a track and the PV. The VELO provides LHCb with a primary vertex resolution of 13 μm in the transverse plane and 71 μm along the beam axis for vertices with 25 tracks, and an impact parameter resolution of less than 35 μm for particles with p_{T} greater than 1 GeV [44].

3.3.2 Magnet

In LHCb, charged particles are bent in the vertical plane by a dipole magnet. Analysis of tracks before and after the particles pass the magnet provides measurement of their momenta through identifying the radius of curvature of the tracks in the magnetic field. Momentum values of up to 200 GeV can be recorded with a resolution of approximately 0.4% using a bending power of 4 Tm [45].

The magnetic field is designed to be uniform in the transverse (x) direction and its polarity is regularly reversed during data taking to take systematic effects of the magnet into account. Approximately half of LHCb's data is recorded at each polarity. A schematic of the magnet can be seen in Figure 3.10.

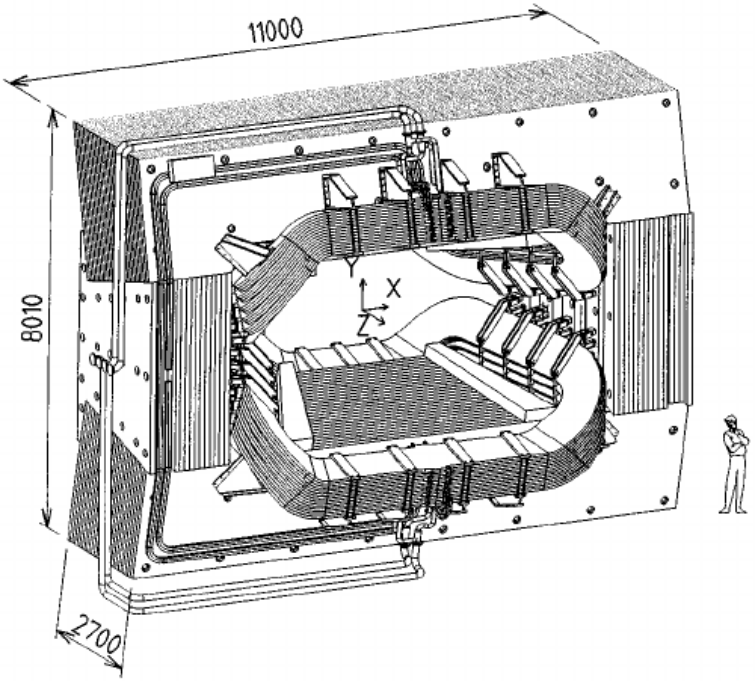

FIGURE 3.10: Schematic of LHCb's dipole magnet [41].

3.3.3 Tracker Turicensis

The Tracker Turicensis (TT) station is positioned between RICH1 and the magnet. It is a 150 cm wide, 130 cm high tracking station that covers the full LHCb acceptance. Four planar silicon detectors are arranged in pairs, in which one is vertical and the other is rotated by an angle of $\pm 5°$ with respect to the vertical axis [46]. These layers are arranged in a (x, u), (v, x) configuration, where x is a vertical layer, u is rotated $+5°$ and v is rotated $-5°$, as shown in Figure 3.11. This arrangement allows for better measurement of the transverse momentum of particles. Each TT layer is comprised of half modules positioned above and below the beampipe, with multiple rows of seven or eight silicon sensors making up each half.

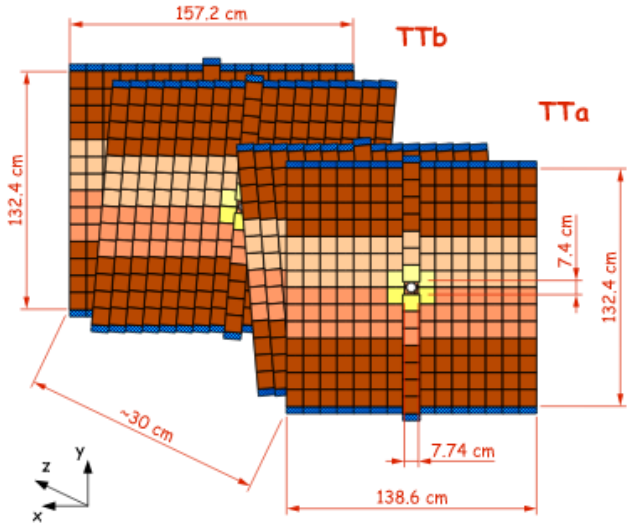

FIGURE 3.11: The (x, u), (v, x) configuration of silicon sensor layers in the TT [46].

3.3.4 Downstream Trackers

The three downstream tracking stations (T1-T3) are comprised of inner and outer tracking regions. In each station, the Inner Tracker (IT) covers a cross-shaped region, approximately 120 cm wide and 40 cm high. This is at the centre of each station where the particle flux is highest and the highest resolution is needed for momentum measurement. The IT is made up of four individual segments, one above, below, to the left and to the right of the beampipe. Each segment is comprised of four layers of silicon microstrip

detectors in the same (x, u), (v, x) configuration as the TT. An example is shown in Figure 3.12.

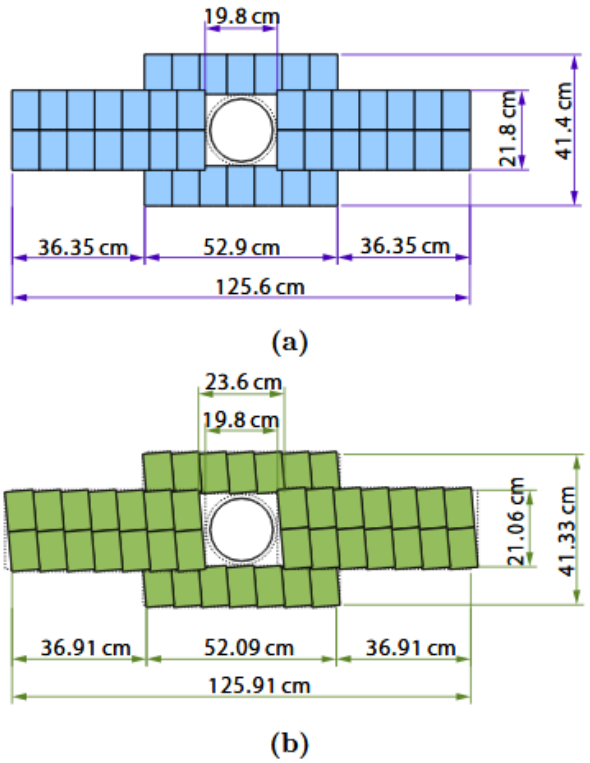

FIGURE 3.12: Layout of an x (top) and u (bottom) layer in the IT with each box representing a sensor [47].

The Outer Tracker (OT) covers the rest of the acceptance in T1-T3, surrounding the IT. The OT is a drift-time tracking detector, built from 5 mm diameter drift-tubes. Each station is made up of four modules arranged in (x, u), (v, x) geometry. A single module consists of two staggered layers of drift-tubes. This can be observed in Figure 3.13.

When charged particles pass through the drift-tubes they ionise the gas molecules inside. Free electrons will then drift towards the centre of the tube to be collected on an anode wire. An electric field applied across the tubes generates an avalanche effect as the electrons drift towards the wire, resulting in an electric current being detected. The time between ionisation and detection is dominated by the drift-time of the ionised electrons and is known to be less than 50 ns. This can be used to determine the co-ordinates of the track passing through the detector.

The drift-tube technology used in the OT has a coarser resolution (approximately 200 μm [48]) than the silicon detectors, but is well suited to covering large areas of the tracking

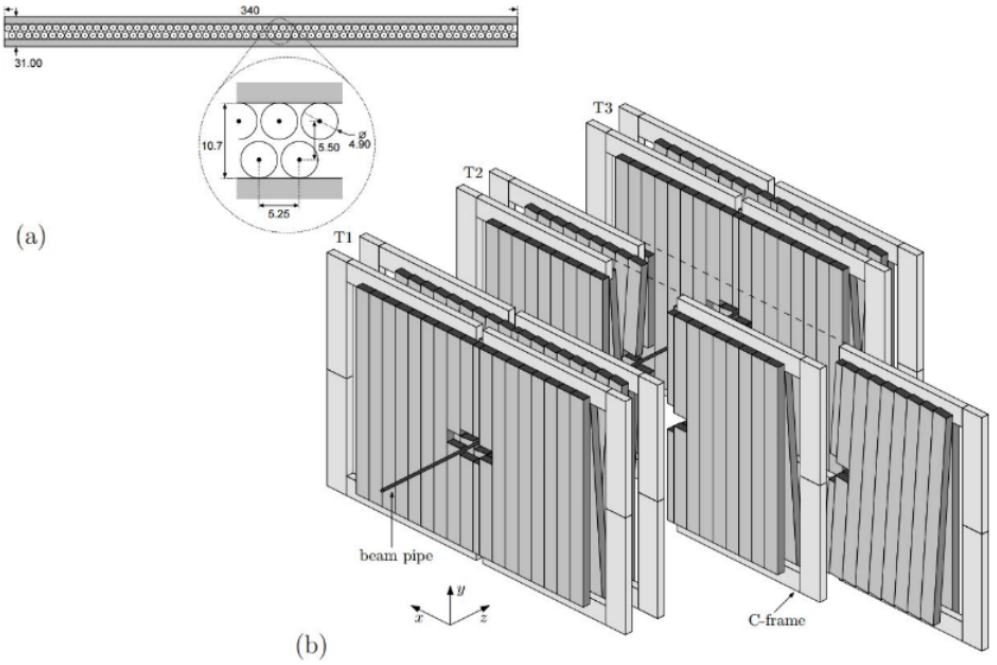

FIGURE 3.13: Schematic of the OT in T1-T3. Cross-section of a detector module with measurements in mm (a) and configuration of the modules (b) [48].

system where particle density is lower.

3.4 RICH

Ring Imaging Cherenkov Detectors (RICH) are used in LHCb's particle identification. The identification is performed through analysis of rings of light produced through the Cherenkov Effect. This occurs when a particle travels through a material faster than light can travel through that medium. The particle will emit a cone of light in the same direction as the track, with a conical angle, θ_C, calculated using

$$\cos\theta_{\mathrm{C}} = \frac{1}{n\beta}, \tag{3.3}$$

where n is the refractive index of the material, and $\beta = v_p/c$, the ratio of the phase velocity of the particle to the speed of light. The Cherenkov angle for a variety of particles is shown in Figure 3.14.

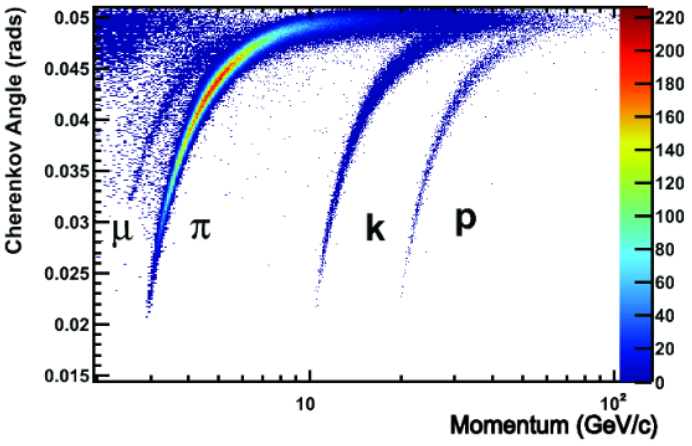

FIGURE 3.14: The reconstructed Cherenkov angle in RICH-1 as a function of track momentum. Distinct curves have been labelled for a variety of particles [49].

The velocity measured for a given particle can be combined with the momentum information obtained for the corresponding track in the tracking system to estimate the particle's mass and so identify it.

The principal purpose of the RICH detectors is to distinguish between pions and kaons for studying B-hadron decays, but the information they obtain is also complementary to the calorimeters and muon chambers for the identification of electrons and muons.

There are two RICH detectors at LHCb optimised for particle identification across a range of momenta [41]. RICH1 is positioned before the magnet and covers the full angular acceptance, it provides K-π separation up to 60 GeV. RICH2 is situated after the downstream tracking stations and distinguishes pions and kaons up to 100 GeV. RICH2 has a slightly reduced angular acceptance of 15 - 120 mrad in the bending plane and 15 - 100 mrad in the non-bending plane. The RICH detectors have an angular resolution of $\sim$ 1 - 3 mrad and can distinguish between pions and kaons at the 3σ level [50].

3.5 Calorimeter System

In LHCb, the calorimeter system is positioned between M1 and M2. Four different stations are used to measure the energy of hadrons, electrons and photons, and to assist

in the identification of these particles. The Scintillating Pad Detector (SPD) and Pre-Shower Detector (PRS) are located before the Electromagnetic Calorimeter (ECAL) and Hadronic Calorimeter (HCAL). A lead converter is placed between the SPD and PRS. The structure of the calorimeter system can be seen in Figure 3.15.

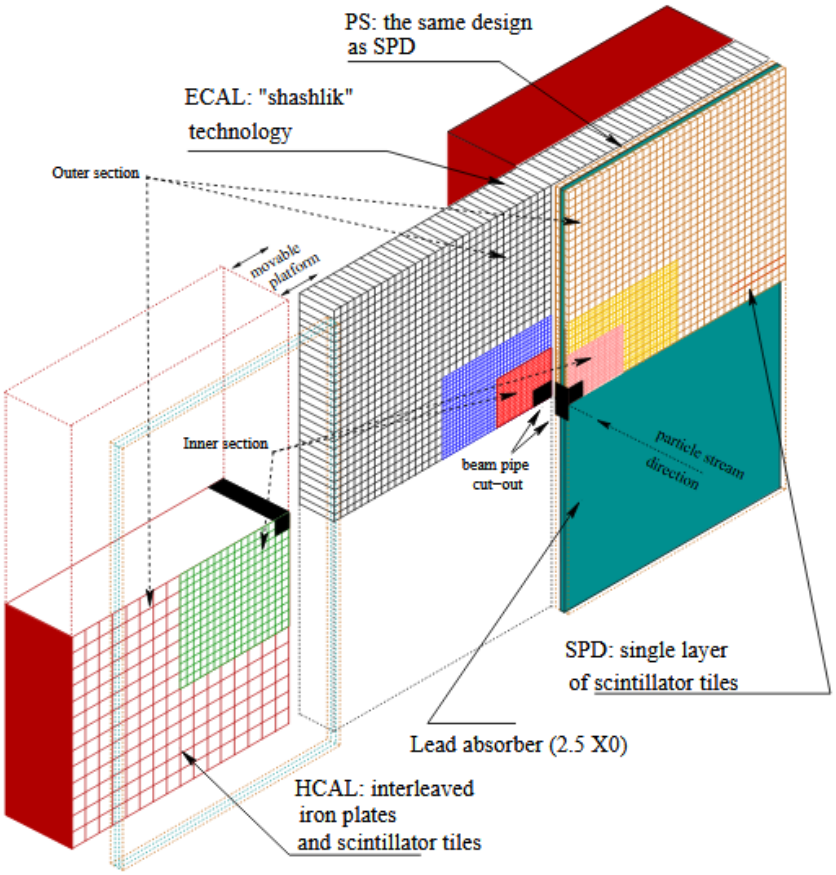

FIGURE 3.15: The calorimeter system [51].

Calorimeters utilise the characteristics of the materials they are built from. When certain particles move through a calorimeter, they produce a cascade of secondary particles known as a shower. Electromagnetic showers are produced by particles that interact primarily or exclusively through the electromagnetic force. Similarly, hadronic showers are produced by particles that interact through the strong force, but they are more complex due to the contribution of different processes to the inelastic production of secondary hadrons.

In the LHCb calorimeter system, particles are identified by the stations in which they deposit their energy. Hits in the SPD can be used to determine whether particles entering the calorimeters are charged or neutral as it is only able to detect charged tracks as no

showering has occurred previously. The lead converter initiates electromagnetic showering which can then be detected in the PRS and used to distinguish between photons and electrons. The SPD/PRS system helps to reject more than 99% of incident pions with an electron retention of above 90% [41]. Both the SPD and PRS consist of scintillating pad detectors.

The ECAL has a shashlik design of 66 layers of alternating scintillating pads and lead absorbers. It is divided into three sections with outwardly increasing cell size to account for the differences in particle density throughout the detector, as shown in Figure 3.16. The energy resolution of the ECAL is given by $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1\%$ [41].

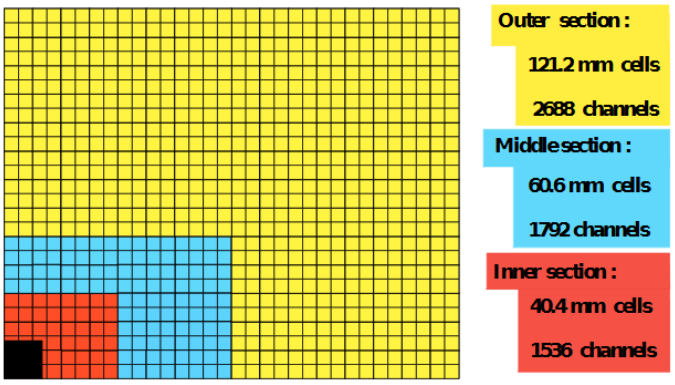

FIGURE 3.16: Cell structure in the ECAL [41].

The HCAL is the most downstream of the calorimeters and is made up of layers of iron and scintillating pads. In contrast to the ECAL, the scintillating material runs parallel to the beam line with each scintillating module interspersed with 1 cm of iron. The HCAL is divided into two sections with outwardly increasing cell size, as shown in Figure 3.17. The energy resolution of the HCAL is given by $\frac{\sigma_E}{E} = \frac{69 \pm 5\%}{\sqrt{E}} \oplus 9 \pm 2\%$ [41].

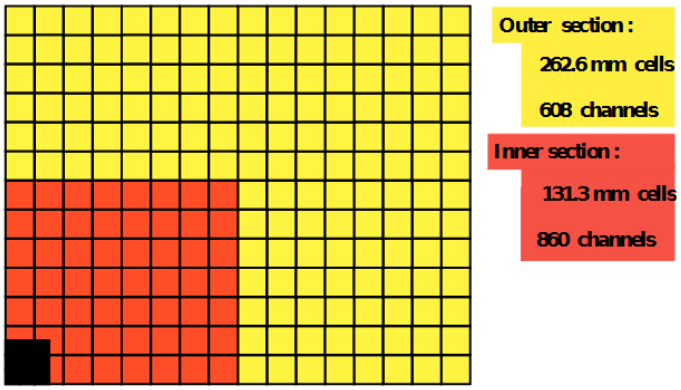

FIGURE 3.17: Cell structure in the HCAL [41].

3.6 Muon System

The muon system consists of five rectangular stations (M1-M5) covering an angular acceptance of ± 300 mrad horizontally and ± 250 mrad vertically [52]. M1 is situated in front of the calorimeter system and M2-M5 are positioned after. Each station is separated into four regions (R1-R4) with the acceptance of each designed to be approximately the same, and the granularity designed to ensure the occupancy is approximately consistent throughout the detector, based on the particle density in each region. M1-M3 have a high spatial resolution in the bending plane as they are used in determining track direction and the transverse momentum of the muons with a resolution of 20% [53].

Stations M2-M5 are preceded by 80 cm thick iron absorbers, which along with the calorimeters, act as muons shields and prevent hadrons and electrons from reaching the chambers. The layout of M2-M5 is is shown in Figure 3.18. A graphical view of the passage of muons through matter can be seen in Figure 3.19.

Multi-Wire Proportional Chambers (MWPCs) are used to detect muons in all five chambers. When a muon passes through one of these MWPCs it ionizes gas molecules and the free electrons are collected on an anode. The innermost region of M1 is different and utilises a Triple-GEM (Gas-Electron-Multiplier) detector. This identifies the muons in a similar way to the MWPCs, but has a higher radiation-hardness which allows for better performance and detector aging in this high particle flux region.

Information from the muon system is used in the trigger, which is discussed in Section 4.4. The momentum resolution of the muon chambers is optimised for muons with momentum of greater than 5 GeV, which is the minimum momentum required to traverse all five stations [41].

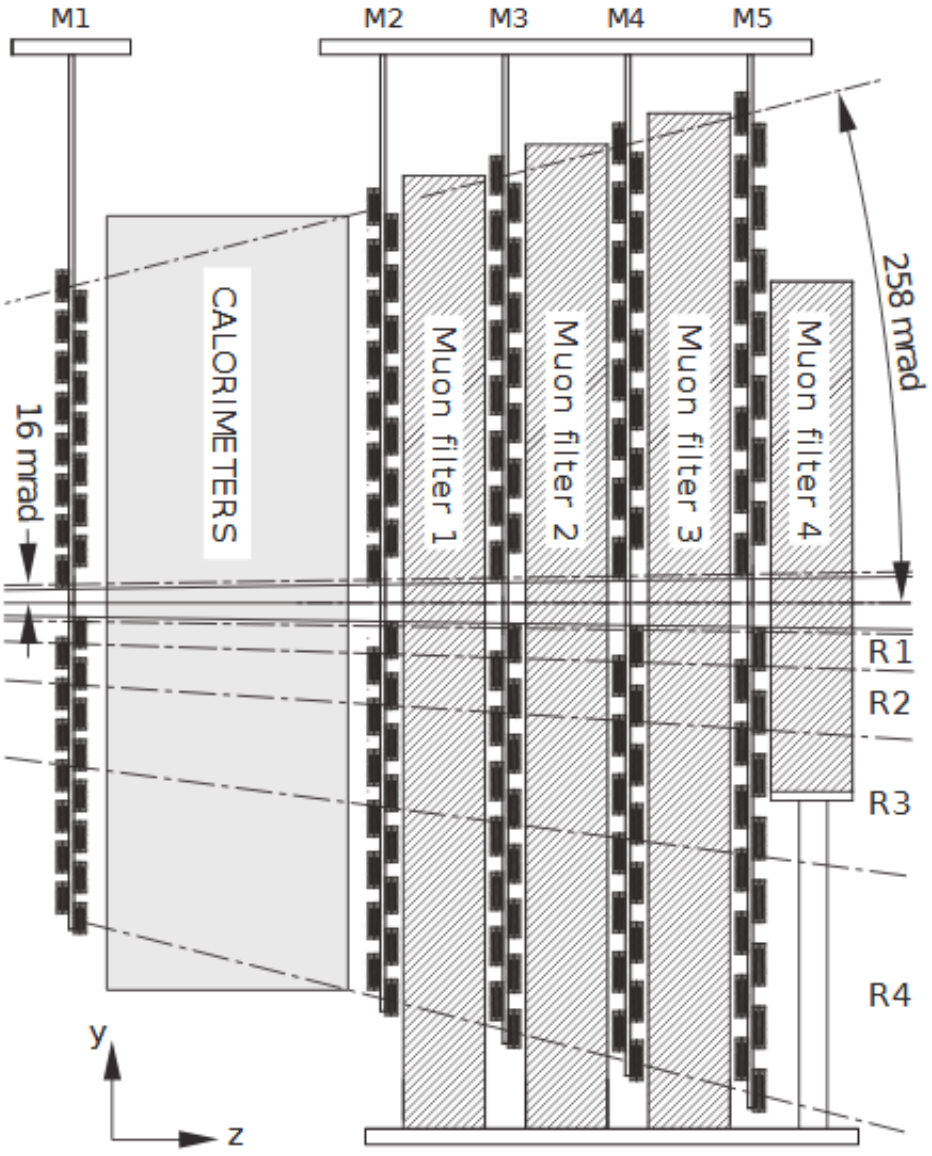

FIGURE 3.18: The layout of the muon stations with the calorimeter system and iron filters in between [41].

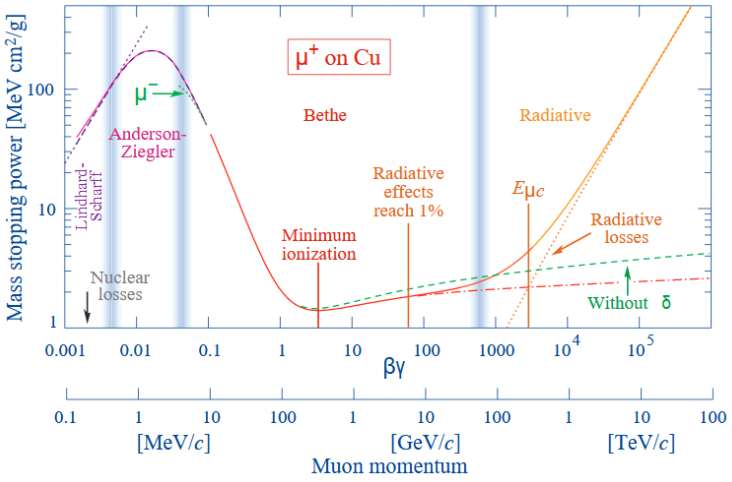

FIGURE 3.19: The passage of muons through matter, showing the mean energy loss per distance travelled [1].

Chapter 4

Event Reconstruction

The particles produced in collisons at LHCb are reconstructed from digitised electrical signals in the detector through a series of online and offline algorithms. Online algorithms are performed in real-time while the collisions are occuring and offline algorithms are run after the data has been written to disk. The online reconstruction is only simple and is used to decide which events should be stored for offline analysis; this is due to the large number of collisions taking place each second and the subsequent time constraints on data processing. The offline reconstruction involves fully reconstructing each event from all of the available information from each sub-detector to enable the precise measurement of final state particles for physics analysis. The following section summarises the offline algorithms for track and vertex reconstruction, and particle identification; the online trigger system used in the first selection of interesting events; and the software chain used to prepare data for physics analysis.

4.1 Track Reconstruction

When charged particles traverse the tracking system they deposit energy as ionisation. This is detected and digitised as a *hit* in each detector the particle interacts with. A trajectory for each particle can be reconstructed by combining these hits along the direction of travel. The momentum of these particles can then be determined by measuring the curvature of bent trajectories and relating this to the strength of the magnetic field. The reconstructed particle paths are known as tracks, with each track being defined by

a set of states tangential to the particle's momentum at various z positions where the sub-detectors make their measurements. A state, $\vec{x}$, is defined as

$$\vec{x} = \begin{pmatrix} x \\ y \\ t_x \\ t_y \\ q/p \end{pmatrix}, \tag{4.1}$$

where x and y are co-ordinates relative to the interaction point at (0,0), with the detector geometry as defined in Section 3.2; t_x and t_y are the slopes of the track in the x and y planes respectively, where $t_x = \frac{dx}{dz}$ and $t_y = \frac{dy}{dz}$; and q/p is the charge divided by the magnitude of the momentum with charge assumed to be $q = \pm 1$.

There are five different categories of tracks defined at LHCb, depending on which sub-detectors have recorded hits contributing to the tracks' reconstruction [54]:

- *Long Tracks* traverse the full tracking system and are the most useful tracks for physics analysis due to the full availability of information and most accurate momentum information. The analyses presented in this thesis use long tracks.
- *VELO Tracks* only have hits in the VELO and therefore have no momentum information but can still be used for primary vertex reconstruction. They can also be used to identify tracks that are travelling in the backwards ($-z$) direction.
- *Upstream tracks* are only detected in the VELO and TT, not the downstream trackers. Typically these are low momentum tracks bent out of acceptance for T1-T3 by the magnet.
- *Downstream Tracks* only leave hits in the TT and downstream trackers, so can be used to reconstruct neutral particles which decay after leaving the VELO or to reconstruct particles outside of the VELO's acceptance.
- *T Tracks* are only detected in the downstream tracking stations. This can occur when, for example, a long-lived neutral particle travels through a significant portion of the detector before decaying.

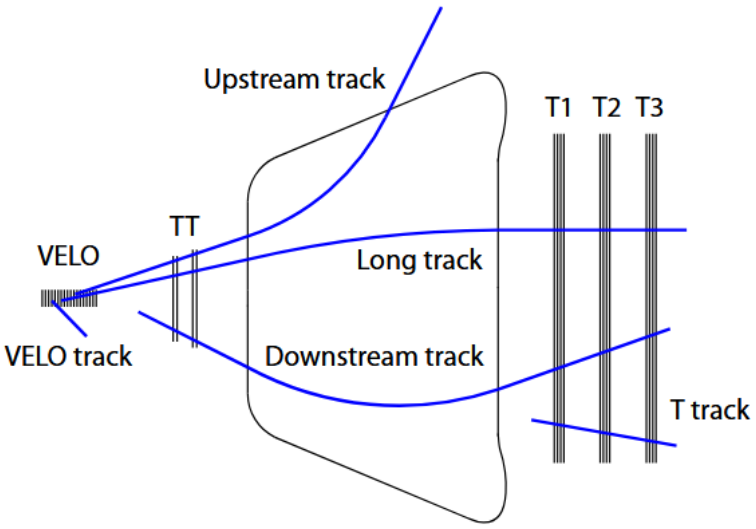

FIGURE 4.1: Different type of tracks [55].

The different types of track are represented graphically in Figure 4.1.

The tracks are reconstructed using pattern recognition procedures to identify measurements belonging to a single track from the other hits present in the event. The track reconstruction begins with VELO Seeding, which searches for straight lines in the VELO. Sets of three hits from modules close to the interaction point are combined with those from the most downstream sensors. A track is then extrapolated backwards from the downstream modules to the interaction point, searching for more measurements close to the path that can be included. The resulting track segments are called VELO seeds and are used in the following algorithms.

T-Seeding follows a similar procedure to VELO seeding [56]. Track segments are formed in the T stations starting with pairs of hits in a given region of the trackers. The resulting segments are known as T seeds.

After the seeding, the track matching process begins. VELO seeds are extrapolated forwards through the detector to single hits in the TT and T stations, then other hits are included in the track if they are close enough to the trajectory. This first step is known as Forward Tracking and approximates the deflection of charged particle trajectories in the magnetic field to a single point-like direction change in the centre of the magnet.

Track Matching then connects unused VELO and T seeds by extrapolating both towards the centre of the magnet and connecting compatible seeds to form Long Tracks. Next,

hits in the TT are matched with the remaining VELO and T seeds to form upstream and downstream tracks, respectively. Any seeds left unmatched after this procedure become VELO or T tracks.

After pattern recognition, track fitting algorithms are used to obtain the tracks. The trajectories mapped by these fits are used to calculate the momentum of the particles that produced the tracks. The fitting procedure is performed using a Kalman filter, which is an iterative procedure used to remove hits within a trajectory and re-fit the trajectory until it matches designated quality requirements [57].

An overall estimate for the efficiency to reconstruct long tracks is found using a tag-and-probe method and $J/\psi \rightarrow \mu^+\mu^-$ decays. Throughout the operation of LHCb, the long track reconstruction efficiency is consistently found to be greater than 95% for long tracks with p_{T} greater than 500 MeV [58]. The momentum resolution for long tracks is ~ 0.5 % for particles below 20 GeV, and $\sim 0.8\%$ for particles around 100 GeV [59].

4.2 Vertexing

The reconstructed point from which particles originate in a collision is known as a vertex. At LHCb, the primary interaction points of the proton beams are known as primary vertices (PVs) and vertices formed by the decays of short-lived particles originating at PVs are called secondary vertices (SVs).

Primary vertex reconstruction is performed using reconstructed tracks to gain a precise knowledge of the origin point of particles [60]. This begins with a seeding process which searches for clusters of tracks alongs the beamline defined by their z co-ordinate, z^{clus}, and the uncertainty associated with is, σ_z^{clus}. The clusters are iteratively merged together, with pairs separated by $D < 5$ merged into a single cluster, where D is defined as

$$D = \frac{\left|z^{clus1} - z^{clus2}\right|}{\sqrt{\left(\sigma_z^{clus1}\right)^2 + \left(\sigma_z^{clus2}\right)^2}}. \tag{4.2}$$

The z^{clu} and σ_z^{clu} of the combined pair is determined using a weighted mean method. The pairing process continues until no more pairs pass the D requirement. The merged clusters produced using this method are then used as seeds in the vertex fitting procedure.

The PV fitting uses the Tukey biweight method [60], a weighted least squares procedure to weight the tracks used to form the vertex seeds based on their χ^2_{ip}. This is defined as the increase of the χ^2 of the PV fit when a given track is included in the PV fitting procedure. The χ^2_{ip} is required to be less than 9. The weight, W_T, of each track is assigned using:

$$\begin{aligned} W_T &= \left(1 - \frac{\chi^2_{ip}}{C^2_T}\right)^2 \quad \text{for } \chi_{ip} > C_T, \\ W_T &= 0 \qquad\qquad\qquad\quad \text{for } \chi_{ip} \leq C_T, \end{aligned} \tag{4.3}$$

where C_T is the Tukey constant [61]. All track types containing hits within the VELO are used in vertex reconstruction.

The position of a PV is then found through the iterative process of minimising χ^2_{PV}, where

$$\chi^2_{PV} = \sum_{i=1}^{n_{tracks}} \chi^2_{ip,i} \cdot W_{T,i}. \tag{4.4}$$

After each iteration the track with the highest impact parameter significance, ip/σ_{ip} is removed, where ip refers to the impact parameter and σ_{ip} is its associated uncertainty. The vertex position is then recalculated with fewer tracks and new impact parameter significances are determined. This process is repeated until all of the tracks in the vertex have an impact parameter significange of less than 4 and the PV candidate is discarded if the final multiplicity is less than 6. The procedure is then repeated for the next seed, but without the tracks that have already been used to reconstruct a PV.

The resolution of the PV can be used as a measure of the performance of the vertex reconstruction procedure. This resolution varies with the number of tracks originating from the PV, from around 10 to 35 μm in the transverse direction and 50 to 250μm in the longitudinal direction [44].

4.3 Particle Identification

Particle Identification (PID) is performed using information from hits in the RICH detectors, calorimeter system and muon chambers. This is done by identifying stable particle

tracks and utilising their interaction properties [41, 62].

The RICH detectors are able to differentiate between particles, particularly pions and kaons, based on their masses, and the calorimeters and muon stations utilise the different interaction properties of photons, electrons, muons and hadrons to identify them.

A basic breakdown of particle identification using the different layers of the detector is shown in Figure 4.2. The identification of photons, electrons, muons and particles that interact with the RICH detectors are discussed below.

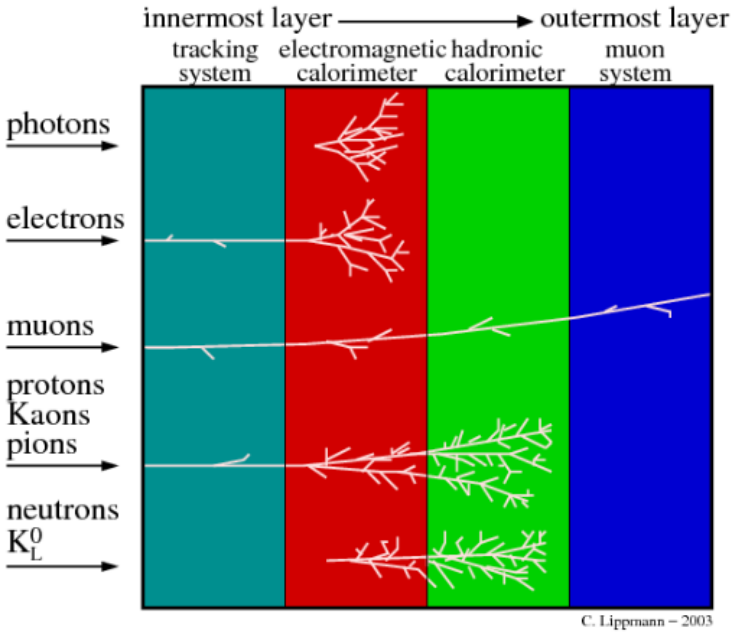

FIGURE 4.2: A graphical representation of the different interactions of particles throughout the layers of a standard particle physics experiment [62].

4.3.1 Photon Identification

Photons are identified by looking for energy deposit clusters in the ECAL where there is no associated track. All charged tracks are extrapolated to the ECAL and a cluster-to-track estimator, χ^2_γ, is evaluated. This gives the proximity of the closest track extrapolation to the cluster. Clusters with $\chi^2_\gamma > 4$ are identified as photons. As discussed in Section 3.5, hits in the SPD/PRS system can also contribute to the separation of photons and electrons. Photons are more likely to shower in the PRS only, while electrons will more likely leave deposits in both the SPD and PRS.

4.3.2 Electron Identification

Electrons are identified using information from the SPD/PRS and ECAL. Similarly to photon identification, a cluster-to-track estimator, χ^2_e, is built to match calorimeter clusters to tracks and corresponding hits in the SPD/PRS. Energy losses through Bremsstrahlung are also taken into account. ECAL clusters corresponding to Bremsstrahlung photons can be predicted and included in the reconstruction of the electron. Further identification criteria is used to identify the high momentum electrons in the following analyses, as described in Section 5.3.2.3.

4.3.3 Muon Identification

Muon identification is performed by extrapolating tracks within the acceptance of M2 and M3 with $p > 3$ GeV into the muon stations. This is the momentum required for a muon to reach M2 and M3, and so tracks with a lower momentum are not considered. A search for hits in the muon stations is performed around a Field of Interest (FOI) at the z position of each muon station along the extrapolated trajectory of the candidate track. The conditions for identifying LHCb's two classes of muons, `isMuon` and `isMuonLoose` are given in Tables 4.1 and 4.2 respectively. For the high energy muons in the following analyses, hits are required in all four of the outermost muon stations.

`isMuon`

Required hits	Momentum range
M2, M3	$p < 6$ GeV
M2, M3, M4 or M5	$6 < p < 10$ GeV
M2, M3, M4, M5	$p > 10$ GeV

TABLE 4.1: Required hits in muon stations M2-M5 for a particle to be fully identified as a muon, satisfying the binary `isMuon` criterion [63].

`isMuonLoose`

Required hits	Momentum range
At least two of M2, M3, M4	$p < 6$ GeV
At least three of M2, M3, M4, M5	$p > 6$ GeV

TABLE 4.2: Required hits in muon stations M2-M5 for a particle to be loosely identified as a muon, satisfying the binary `isMuonLoose` criterion [63].

The identification efficiency for muons is measured to be above 94%. There is a 3% misidentification rate for pions misidentified as muons which drops to 1% when PID performed by the RICH detectors is combined with that performed by the muon stations [41].

4.3.4 RICH Identification

Hit patterns in the RICH detectors are used to identify particles with associated reconstructed tracks. The patterns are compared to the patterns expected for Cherenkov angles for particles with different mass hypotheses and a likelihood variable is constructed for each hypothesis. The momentum coverage of the RICH detectors does not extend beyond 150 GeV and so it cannot in general be used to distinguish between high momentum particles, for example produced by decays of W and Z bosons.

4.3.5 Jet Reconstruction

Jets are showers of particles produced during the hadronisation of hard-scattering final state partons, such as quarks. At LHCb, jets are reconstructed using the standard particle flow algorithm [64].

The particle flow algorithm is designed to use information from the tracking system, calorimeters and PID detectors to provide a list of input particles that can then be reconstructed into jets. The four momenta of individual tracks from a given primary vertex are combined with deposits in the calorimeters and other PID information to produce this list. A jet clustering algorithm is then used to group particles into jets using the anti-k_t method [65] with a radius $R = \sqrt{\delta\phi^2 + \delta\eta^2} = 0.5$.

It is ensured that only high quality jets are selected for use in physics analysis by applying the following requirements [66]:

- At least two particles in the jet must point back to the primary vertex„
- No one charged particle should carry more than 60% of the jet p_{T},
- The jet must contain a charged particle with p_{T} greater than 1.2 GeV,
- The fraction of the jet made up of charged particles must be at least 10%.

The efficiency with which jets are reconstructed in the following analysis is discussed in detail in Section 5.5.1.6.

4.4 Triggers

The LHCb trigger is a hardware and software system designed to decide whether events are interesting enough to be written to disk for offline analysis [67]. This is necessary as the collision rate in LHCb is very high at around 40 MHz and exceeds the limitations of data storage capacity and I/O speeds. The trigger system consists of the Level-0 trigger (L0) and two Higher Level Triggers (HLT1 and HLT2). L0 is a hardware trigger, and HLT1 and HLT2 are software triggers. Only events that pass all three stages of the trigger are recorded.

The three trigger stages are outlined in the following section. The Run-II trigger scheme is summarised in Figure 4.3.

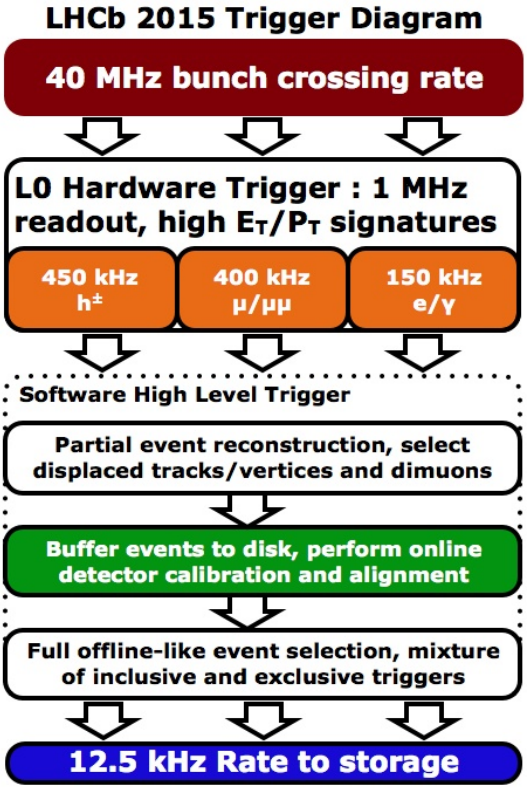

FIGURE 4.3: A summary of the Run-II trigger scheme with data outputs at each level displayed [68].

4.4.1 L0

The L0 hardware trigger is the first stage of event selection at LHCb. L0 is split into two categories - calorimeter-based triggers and muon triggers - followed by the L0 Decision Unit which uses the trigger information to decide whether a given event should be passed on to HLT1 or discarded in less than 4 μs [69]. Through this method, L0 is used to reduce the data rate from 40 MHz to approximately 1 MHz.

4.4.1.1 Calorimeter Trigger

The L0 calorimeter triggers use information from the calorimeter system (described in Section 3.5) to identify the presence of a hadron, electron or photon with high transverse energy, E_{T}. First, the ECAL and HCAL select high-E_{T} deposits - $E_{\mathrm{T}} > 3.7$ GeV for hadrons and $E_{\mathrm{T}} > 3$ GeV for photons and electrons [70]. The information from the selected clusters is combined with information from the SPD/PRS to determine whether the clusters were deposited by a hadron, electron or photon. Only the candidates with the highest transverse energy are used in the trigger decision. Additionally, the SPD multiplicity, n_{SPD}, is recorded to be used by the L0 Decision Unit to discard events with the highest track multiplicities, as these would take up a disproportionately large amount of storage space and processing time.

4.4.1.2 Muon Trigger

The L0 muon trigger is controlled by four processing units, one for each transverse quadrant of the muon stations. For each event, each processing unit selects the two muon candidates with the highest transverse momentum, p_{T}. The high-p_{T} muon candidates are required to have left a straight track segment through the muon stations, which also points back to the interaction point. The track is extrapolated from hits in M3 to M2, M4 and M5. If at least one matching hit is found, the extrapolation is extended to M1 and the information from M1 and M2 is then used to calculate an estimate of the track p_{T}. The information for each of the selected candidates is passed on to the L0 Decision Unit.

4.4.1.3 L0 Decision

The L0 Decision Unit gathers the information from the calorimeter and muon triggers to decide which events should be sent on to HLT1. Events are selected if they satisfy a number of different criteria, with thresholds chosen to select events containing high-p_{T} muons and electrons, hadrons, photons or neutral pion candidates.

The L0 trigger lines used in the following analyses are summarized in Table 4.3.

L0 Trigger Lines

Line	Thresholds
`L0Muon`	$n_{SPD} < 450, p_{\mathrm{T}} > 2.8$ GeV
`L0EWMuon`	$n_{SPD} < 10000, p_{\mathrm{T}} > 6.0$ GeV
`L0Electron`	$n_{SPD} < 450, p_{\mathrm{T}} > 2.8$ GeV

TABLE 4.3: Examples of Run-II L0 trigger thesholds with n_{SPD} referring to the multiplicity at the SPD sub-detector.

4.4.2 HLT1

HLT1 is the first software stage of the trigger. Its purpose is to accept events from L0 and validate or discard decisions to reduce the output rate to 150 kHz, ready to be passed on to HLT2. Similarly to L0, the event selections are performed using a set of trigger lines, with the lines used in the following analyses summarised in Table 4.4. The decision time per event at HLT1 is 35 ms [70].

HLT1 makes decisions by first accepting events from L0 and performing a simplified track fitting procedure, where T seeds are reconstructed with an assumption that they originate from the interaction point. Next, simplified 2D VELO tracks are reconstructed in the $r-z$ plane and compared with a track χ^2 criterion. If they pass this and originate from the interaction point, ϕ sensor information is included so 3D VELO tracks can be reconstructed. The 2D tracks are used to reconstruct primary vertices, and then the VELO and T station information is combined and used to accept or reject each event.

HLT1 Trigger Lines

Line	Selection requirements

`HLT1SingleMuonHighPT`	`L0Req = L0Muon or L0MuonEW`
	$p > 6.0$ GeV
	$p_T > 4.34$ GeV
	`TrChi2` < 3
	`MaxOTHits = 15000`
	`MaxITHits = 3000`
	`MaxVeloHits = 6000`

`Hlt1SingleElectronNoIP`	`L0Req = L0Electron`
	$p > 20.0$ GeV
	$p_T > 10.0$ GeV
	`TrChi2` < 3
	`MaxOTHits = 15000`
	`MaxITHits = 3000`
	`MaxVeloHits = 6000`

TABLE 4.4: Examples of Run-II HLT1 trigger lines. The `L0Req` requirement is used to select events passing the given L0 line. The `TrChi2` variable allows for quality cuts on track fits. The `MaxOTHits`, `MaxITHits` and `MaxVeloHits` requirements are Global Event Cuts limiting the maximum multiplicity of the each event in the OT, IT and VELO respectively.

4.4.3 HLT2

The full offline event reconstruction is performed by HLT2. The event selection is performed and the output data is cut to 12.5 kHz. As for the previous stages, the event selection is performed using a set of trigger lines, with those used in the following analyses summarised in Table 4.5. These are more complex than for HLT1 and can be designed for different physics processes with the aim of reducing the output to a level that can be written to storage for use in offline analyses. HLT2 therefore has a longer decision time than HLT1 of 650 ms per event [70]. After an event passes HLT2 it is written to storage with the total size of a raw event being $\approx$ 31 kB [67].

4.4.4 Real-time Alignment and Calibration

At the beginning of Run-II, LHCb became the first LHC collider experiment to employ a novel method of real-time online alignment and calibration. This is performed at the start of every LHC fill or more often if necessary. Dedicated samples are collected by

HLT2 Trigger Lines

Line	Selection requirements

`HLT2EWSingleMuonVHighPt`	`Input = BiKalmanFittedMuons`
	`L0Req = L0Muon or L0MuonEW`
	`Hlt1Req = MuonDecision`
	$p_T > 12.5$ GeV

`Hlt2EWSingleElectronVHighPt`	`Input = BiKalmanFittedElectrons`
	`L0Req = L0Electron`
	`Hlt1Req = ElectronDecision`
	$p_T > 1.5$ GeV
	`PrsMin = 50 MeV`
	`EcalMin = 0.1`
	`HCALMax = 0.05`
	`TkChi2` < 20

TABLE 4.5: Examples of Run-II HLT2 trigger lines. The `Input` is the location from which the trigger line should select the required particle candidates. The `L0Req` and `Hlt1Req` are used to only select events that have passed the given L0 and HLT1 lines respectively. The `PrsMin`, `EcalMin` and `HCALMax` variables define the minimum or maximum deposits required in the PRS, ECAL and HCAL respectively. `TkChi2` defines the cut on the minimum quality of the track fit.

HLT1 and used to calculate new calibration constants which are employed at HLT2 to improve data quality [71].

4.5 Software

The LHCb software is built on the GAUDI [72] framework and is used to produce the output data used in offline analysis. It contains a number of different software applications used in simulation, reconstruction and analysis.

Simulated Monte Carlo (MC) events are used to model the interactions of specified decays within the detector to gain a better understanding of its response and performance. This begins with generating simulated pp interactions to produce the desired particles and their decays. In the following analyses, MC events are generated at LO with PYTHIA [73], which is able to simulate large numbers of events using a variety of PDF sets. Further information on the simulated samples used in the following analyses can be found in Sections 5.2 and 6.2.

At LHCb, the GAUSS [74] project is used to configure the Monte Carlo generators. It also configures the simulation of the interaction of the simulated particles with the

detector material and the physical response of the detector using GEANT4 [75]. The BOOLE application [76] continues the process by simulating the readout of the detector and digitising the simulated data so that it becomes indistinguishable from real data. The BRUNEL application [77] is then used to perform the reconstruction of the digital data, either for simulation from BOOLE or for real data from the detector. The software project used in the trigger is called MOORE [78].

The data needed for physics analysis is processed and retrieved using a global computing network known as the GRID [79]. The submission of the code used to run over this data on the GRID is handled by DIRAC [80] and GANGA [81]. The final physics analysis is predominantly performed by individual users using the DAVINCI [82] framework.

An overview of the software chain used to process data is shown in Figure 4.4.

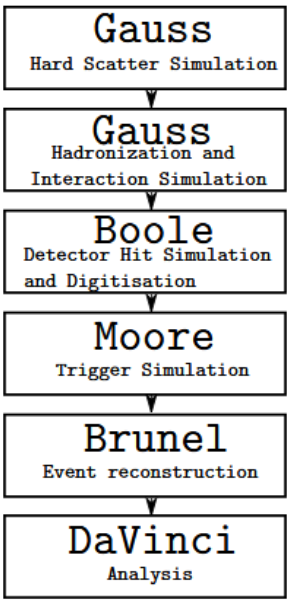

FIGURE 4.4: A summary of LHCb software chain. [83].

Chapter 5

Top Quark Pair Production Cross-section

This chapter describes the primary analysis presented in this thesis: the measurement of the top quark pair production cross-section through the partial reconstruction of the μeb final state. The motivation for making such a measurement at LHCb is outlined first, followed by a description of the methods used to select events, determine the background expectation and evaluate the efficiency with which the events are reconstructed and selected. The value for the cross-section is then calculated along with the associated uncertainties, and the final result is presented and compared with theoretical predictions.

5.1 Motivation

As a forward region experiment, LHCb is able to observe top quark production in a unique phase space which is inaccessible to the central detectors. This can be seen in Figure 5.1. For final state particles to be detected by LHCb, the interactions that produce them typically occur between one high and one low Bjorken-x parton, as shown by Figure 5.2. These interactions produce a high enough momentum transfer to push the particles into the LHCb acceptance.

Due to LHCb's limited acceptance, partial reconstruction of the final state is the most favourable option for measuring the top quark pair production cross-section. A number

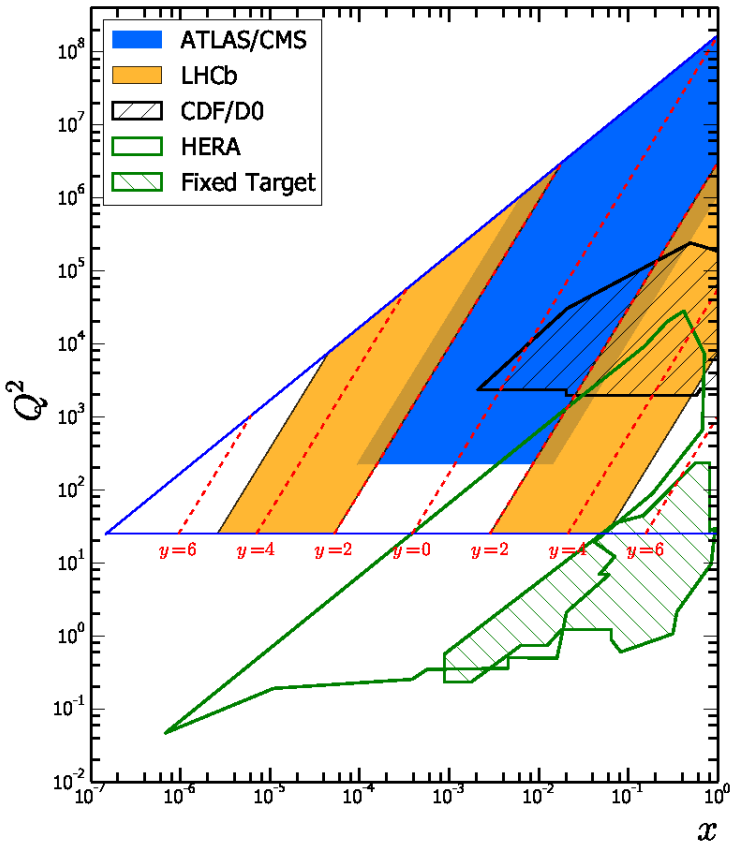

FIGURE 5.1: The kinematics of LHC experiments at 13 TeV [84].

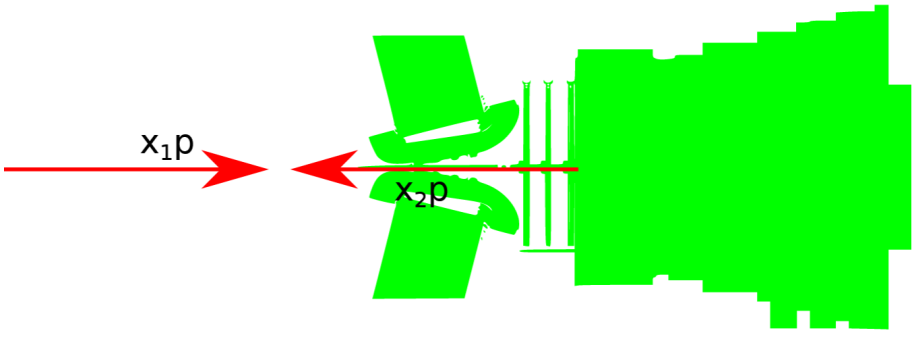

FIGURE 5.2: Graphical representation of a pp collision at LHCb. For the particles produced in the interaction to enter LHCb acceptance, parton 1 must have a high Bjorken-x while the Bjorken-x of parton 2 must be low for the momentum transfer to boost the final state particles into the forward region [84].

of possible final states can be identified in the Feynman diagram shown in Figure 5.3 and are summarised in Table 5.1.

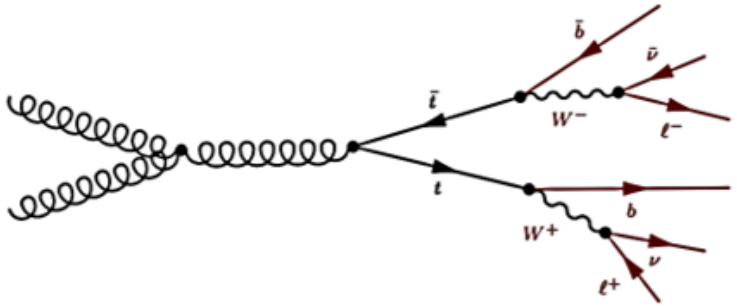

FIGURE 5.3: Feynman diagram of top quark pair production through gluon fusion decaying to leptons (ℓ) and neutrinos (ν).

Run-II is an exciting time for investigating top physics at LHCb. The increase in centre-of-mass energy has increased the expected inclusive $t\bar{t}$ cross-section by a factor of ∼10, as can be seen in Table 5.1. Therefore, final states that were inaccessible in Run-I can now be explored.

$d\sigma$ (fb)	7 TeV	8 TeV	14 TeV
lb	285 ± 52	504 ± 94	4366 ± 663
lbj	97 ± 21	198 ± 35	2335 ± 323
lbb	32 ± 6	65 ± 12	870 ± 116
$lbbj$	10 ± 2	26 ± 4	487 ± 76
l^+l^-	44 ± 8	79 ± 15	635 ± 109
l^+l^-b	19 ± 4	39 ± 8	417 ± 79

TABLE 5.1: The predicted production cross-sections of $t\bar{t}$ final states at 7, 8 and 14 TeV within LHCb acceptance using POWHEG predictions matched to PYTHIA8. The uncertainties include contributions from scale, PDF, showering and jet tagging [85].

In general, the final states with the highest yield are those that require fewer objects to be reconstructed. This is because the production of the $t\bar{t}$ system requires such a large momentum transfer that it is rare for a collision to result in enough forward momentum to boost all of the observable final state particles into the detector acceptance. The fewer objects required, the more events are visible to LHCb. As shown in Table 5.1, the llb final state is expected to produce the lowest yield in Run-II. However, it it also expected to have the highest purity.

Purity increases with the number of objects required in the final state. In Run-II, the signal-to-background ratio has also increased due to a combination of the increase in

centre-of-mass-energy and the partial reconstruction of final states. The cross-sections of backgrounds produced at lower Bjorken-x than the complete $t\bar{t}$ system have not increased as much as the $t\bar{t}$ cross-section, meaning these backgrounds are supressed.

As discussed in Section 2.3, top quark production at LHCb is dominated by gluon-gluon fusion. Measuring the $t\bar{t}$ production cross-section therefore provides significant constraints on gluon PDFs at high Bjorken-x, where large uncertainties are currently present. Due to the final states only being partially reconstructed, different final states probe different Bjorken-x depending on the number of objects observed by the detector and the momentum required to boost them into the forward acceptance. The mean Bjorken-x probed in each final state is given in Table 5.2.

Final state	Mean Bjorken-x
lb	0.295
lbb	0.368
μeb	0.348
μebb	0.415

TABLE 5.2: The mean Bjorken-x of the most energetic parton probed by different partially reconstructed $t\bar{t}$ final states [86].

A measurement of the $t\bar{t}$ production cross-section at LHCb with the same ~4% uncertainty achieved by ATLAS/CMS in Reference [87] would allow for a 20% reduction in the gluon PDF uncertainty as can be seen in Figure 5.4.

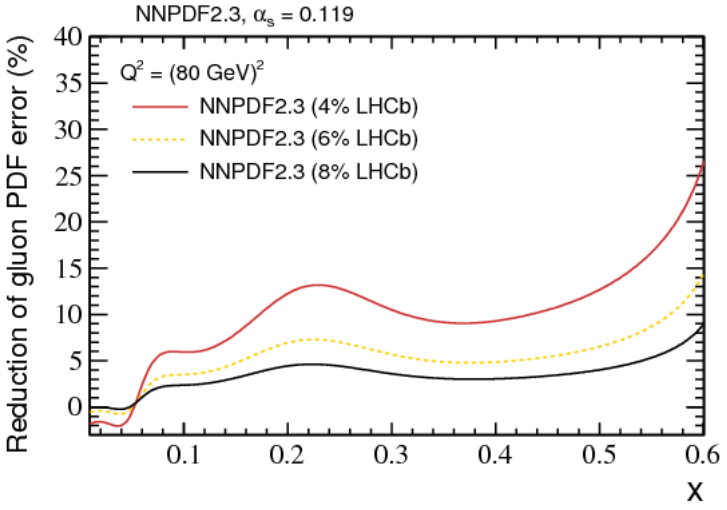

FIGURE 5.4: Predicted reduction in gluon PDF uncertainty with the inclusion of an LHCb $t\bar{t}$ cross-section measurement with uncertainty of 4-8% [27].

In Run-II, the μeb channel offers the highest purity final state. Requiring both leptons supresses backgrounds from $W + b\bar{b}$ and multijet production, and choosing different flavour leptons supresses $Z + b\bar{b}$. This channel was out of statistical reach in Run-I, but the measurement is now possible with more statistics from the increased centre-of-mass energy.

5.2 Data and Simulation

This analysis is performed using the full LHCb Run-II dataset, collected at $\sqrt{s} = 13$ TeV. This corresponds to an integrated luminosity of 5.4 fb^{-1}. A breakdown by year can be seen in Table 5.3.

At the time of writing, the luminosity calibration for 2017 and 2018 data sets is unavailable. To determine a value for the recorded integrated luminosity in these years, an investigation of $Z \to \mu\mu$ data is performed. The number of $Z \to \mu\mu$ events produced in each year is compared with the number produced in 2015 and 2016 combined, which has a known luminosity of 1.96±0.08 fb^{-1}. An additional uncertainty of 2% is then added in quadrature to account for any changes in the $Z \to \mu\mu$ reconstruction efficiency between years.

Year	Integrated Luminosity (fb^{-1})
2015	0.29±0.01
2016	1.66±0.06
2017	1.58±0.07
2018	1.86±0.08
Total	5.40 ± 0.24

TABLE 5.3: The integrated luminosity recorded by LHCb for each year in Run-II as used in this analysis.

The simulated samples used in this analysis are produced at LO as discussed in Section 4.5 using PYTHIA8 and the CT09MCS[88] PDF set. Two separate $t\bar{t}$ MC samples are produced: one where the top quarks are produced through gluon-gluon fusion and the other through $q\bar{q}$ annihilation. The samples are weighted by their LO cross-sections and combined to form a single $t\bar{t}$ sample. NLO corrections for the MC samples are calculated using AMC@NLO [89] and POWHEG [90, 91] with the NNPDF3.0[92] PDF set.

5.3 Event Selection

The following section describes the event selection applied to isolate the $t\bar{t}$ signal and reduce contributions from the backgrounds described in Section 5.4. The event triggers, selection criteria applied to the leptons and requirements placed on the jet are outlined below. The selection builds on previous studies such as Reference [93].

5.3.1 Trigger Requirements

Before being assessed against the full selection criteria, candidate events must first pass all three stages of the LHCb trigger, as described in Section 4.4. For this analysis, events are triggered by the muon, which is required to pass the `L0EWMuon` line at the hardware level, and the `HLT1SingleMuonHighPT` and `HLT2EWSingleMuonVHighPt` lines at the software level.

5.3.2 Lepton Selection

Both leptons are required to have p_{T} greater than 20 GeV and to be inside the LHCb acceptance ($2.0 < \eta < 4.5$). The leptons must also have opposite charge, i.e. either $e^{+}\mu^{-}$ or $e^{-}\mu^{+}$. Further requirements are placed on the impact parameter, isolation and identification of the leptons as described below.

5.3.2.1 Impact Parameter

The impact parameter, as defined in Section 3.3.1, of each lepton must be less than 0.04 mm to remove events in which the leptons are not prompt. For example, tau leptons produced in the PV will travel a finite distance before decaying and so the decay products will have a larger impact parameter. Therefore, placing an *ip* requirement on the signal data will suppress backgrounds involving a $Z \to \tau\tau$ decay with a muon and electron pair in the final state. The motivation for this requirement can be observed in Figure 5.5 in which the impact parameter distributions for $t\bar{t}$ and $Z \to \tau\tau \to \mu e$ MC are compared. The events selected for this plot are chosen with a minimal selection, requiring each lepton to have a p_{T} greater than 20 GeV and $2.0 < \eta < 4.5$. Placing the requirement at 0.04 mm eliminates ∼60% of the $Z \to \tau\tau$ events and retains ∼75%

of the $t\bar{t}$ signal. However, candidate events are still expected to contain a contribution from $Z \rightarrow \tau\tau$ + jet contamination, described in Section 5.4.2.

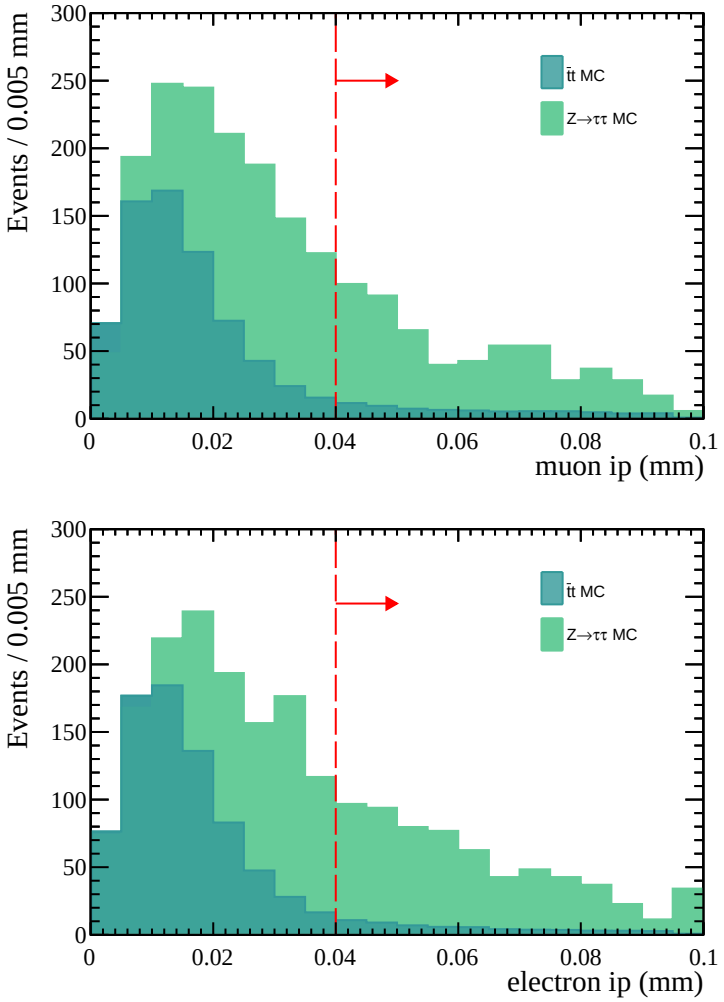

FIGURE 5.5: An overlay of impact parameter distributions to compare $t\bar{t}$ and $Z \rightarrow \tau\tau \rightarrow \mu e$ MC for the muon (top) and electron (bottom). The distributions are normalised to 5.4 fb^{-1}. The dashed red lines shows the position of the *ip* requirement with the arrow indicating which events are rejected.

5.3.2.2 Isolation

The leptons are also required to be isolated, to reduce the potential lepton misidentification background of particles produced in hadronic jets being identified as muons. Lepton misidentification contributes a background when one or both final state particles in the event are hadrons misidentified as leptons. Such misidentification can arise from multijet events, W or Z boson production in association with a jet, or from other $t\bar{t}$ final states (such as semi-leptonic decay). Misidentification can occur when pions and kaons are

tagged as electrons after leaving large deposits in the ECAL, or are tagged as muons when they punch-through to the muon chambers or decay to muons in-flight.

To reduce the contribution from this background, an isolation variable, $p_{\mathrm{T}}^{\mathrm{Cone}}$, is defined and a requirement is placed on each lepton. The $p_{\mathrm{T}}^{\mathrm{Cone}}$ of the lepton is the sum of the transverse momentum in a cone of radius $R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.5$ around the track:

$$p_{\mathrm{T}}^{\mathrm{Cone}} = \left(\sum_{i}^{N} p_i \right)_T \tag{5.1}$$

where N is the number of additional tracks counted in the cone. Signal leptons produced in the decay will have very little activity around their tracks in comparison to particles produced in jets. Therefore, requiring the leptons to be isolated (i.e. requiring they have a low $p_{\mathrm{T}}^{\mathrm{Cone}}$) reduces the contribution from the lepon misidentification background.

Isolation distributions for opposite- and same-sign electron-muon pairs in data are compared in Figure 5.6. Due to charge imbalance, same-sign pairs cannot be produced in $t\bar{t}$ events and are considered a control region for the lepton misidentification background as discussed further in Section 5.4.1. The events used to produce the plots in Figure 5.6 are selected by requiring the leptons to have p_{T} greater than 20 GeV and to be within LHCb acceptance. The muon is also required to pass all three stages of the trigger.

In the $t\bar{t}$ selection, $p_{\mathrm{T}}^{\mathrm{Cone}}$ is required to be less than 5 GeV. This requirement removes ∼90% of the same-sign events and so significantly reduces the lepton misidentifiction background. This background accounts for ∼70% of the signal sample before the requirement is applied, and is reduced to ∼30% after.

In related analses, such as [94], a tighter requirement is placed at 2 GeV to increase the signal purity, but in this analysis a slightly relaxed requirement is chosen due to the low statistics of the sample. In future measurements with an increased data set where statistical uncertainty will not be a problem, the requirement may be tightened again.

5.3.2.3 Identification Requirements

As discussed in Section 4.3.3, a requirement is placed on all muons to ensure that they have been properly identified. This standard binary selection is called `isMuon`. For a

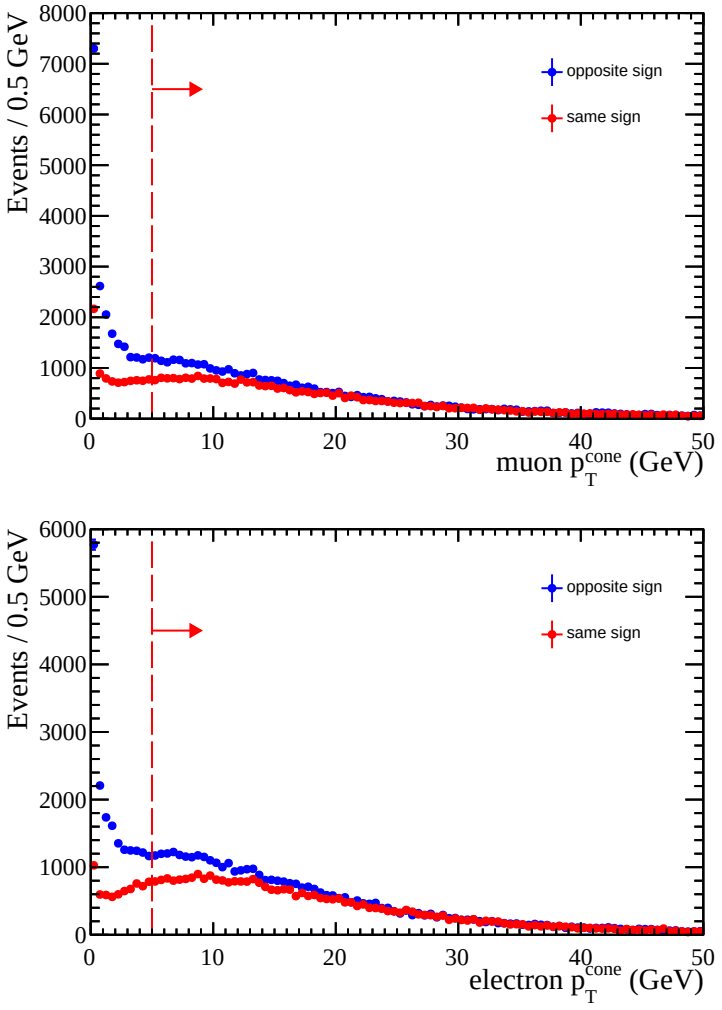

FIGURE 5.6: Comparisons of p_T^{Cone} distributions in opposite- and same-sign data in blue and red respectively, for muons (top) and electrons (bottom). The red dashed line shows the position of the p_T^{Cone} requirement with the arrow showing the region being rejected.

high p_T muon to pass the `isMuon` requirement, it must have associated hits in all four of the outermost muon chambers.

In addition to the basic requirements used to identify electrons in Section 4.3.2, the high p_T electrons in this analysis are assessed against further criteria based on the selection in Reference [95]. The electrons are required to deposit more than 50 MeV in the PRS, more than 10% of their energy in the ECAL and less than 5% of their energy in the HCAL. The electron is also required to fail the binary `isMuonLoose` requirement, which requires at least three hits in the four outermost muon chambers, as defined in Section 4.3.3. In later sections, these four requirements are referred to as the electron identification criteria.

5.3.3 Jet Selection

Jets are selected using the standard jet reconstruction described in Section 4.3.5. Additionally, the jet must have a p_{T} greater than 20 GeV and be separated from each lepton by a radius of 0.5 in $\eta - \phi$ space. A reduced pseudorapidity range of $2.2 < \eta < 4.2$ is used to ensure a flat jet reconstruction and identification efficiency.

The jet must also have been tagged by the secondary vertex (SV) tagger as outlined in Reference [96]. First, multiple two-body vertices are constructed using pairs of displaced tracks. Next, any two-track vertices that share tracks are linked together to form an n-body SV. Then a number of kinematic and quality requirements are placed on the vertices, some of which are discussed in Section 5.5.2. A jet is SV-tagged if it contains an SV within $\Delta R < 0.5$ of the primary jet axis.

5.3.4 Summary

The event selection described in the text is summarised in Table 5.4. After the requirements are applied to the Run-II dataset, 132 $t\bar{t}$ candidates are selected. Expected background contributions to this sample are discussed in Section 5.4.

Trigger	`L0EWMuon`
	`HLT1SingleMuonHighPT`
	`HLT2EWSingleMuonVHighPt`

Muon	`isMuon` $p_{\mathrm{T}} > 20$ GeV, $2.0 < \eta < 4.5$ $ip < 0.04$ mm, $p_{\mathrm{T}}^{\mathrm{Cone}} < 5$ GeV
Electron	$E_{PRS} > 50$ MeV, $E_{HCAL}/P < 0.05$ $E_{ECAL}/P > 0.1$, `!isMuonLoose` $p_{\mathrm{T}} > 20$ GeV, $2.0 < \eta < 4.5$ $ip < 0.04$ mm, $p_{\mathrm{T}}^{\mathrm{Cone}} < 5$ GeV
Leptons	$\Delta R(\eta, e) > 0.1$
Jet	$p_{\mathrm{T}} > 20$ GeV, $2.2 < \eta < 4.2$, $\Delta R(l, j) > 0.5$ SV-tagged

TABLE 5.4: A summary of the selection criteria applied to isolate $t\bar{t}$ signal.

5.4 Signal Purity

When the event selection described in Section 5.3 is applied, 132 $t\bar{t}$ candidates are selected. Some of these candidates are contributions from background processes. The sources of background considered are lepton misidentification, Z + jet, single top production, di-boson production and misidentified jets. Contributions from each source will be discussed in greater detail below.

5.4.1 Lepton Misidentification

The lepton misidentification background is modelled by considering same-sign events using $e^+\mu^+$ and $e^-\mu^-$ final states. These lepton pairs cannot have been produced in $t\bar{t}$ signal events due to the charge imbalance of the leptons and so are used to determine the likelihood of lepton misidentification occurring in the signal data. These pairs typically consist of one fully reconstructed lepton and a misidentified pion arising from $(W \rightarrow \mu\nu)$ + jet, Z + jet or QCD events.

In this method, a signal region and three control regions are defined. To estimate the contribution in the signal region, Region A, events are selected in three control regions: B, C and D.

In Region B, the selection criteria is the same as in Section 5.3, but the leptons are required to have the same charge. In Region C, opposite-sign events are selected with the electron identification requirements defined in Section 5.3.2.3 reversed. In Region D, same-sign events are selected with the same anti-electron identification criteria as for Region C. In Regions C and D the SV-tagging requirement applied to the jet in the event is removed to achieve a higher statistical precision.

A correction factor is then calculated by dividing the number of events selected in Region C by the number of events selected in Region D. This correction factor is applied to the number of events selected in Region B to estimate the expected contribution in Region A. This can be mathematically represented as:

$$N^{\mathrm{A}} = N^{\mathrm{B}} \cdot (N^{\mathrm{C}}/N^{\mathrm{D}}) \tag{5.2}$$

where N^{A} is the number of events expected in Region A and N^{B}, N^{C} and N^{D} are the number of events selected in Regions B, C and D respectively.

A graphical representation of the distinctions between the four regions is shown in Figure 5.7 for further clarity.

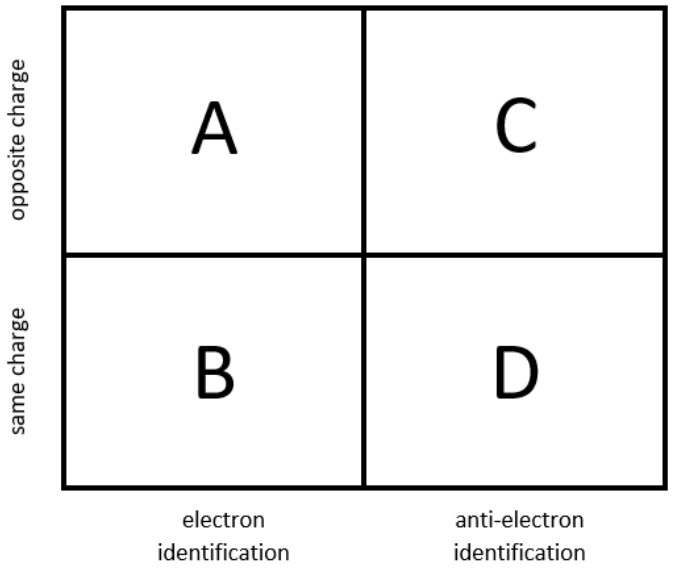

FIGURE 5.7: A graphical representation of the signal region, A, and three control regions, B, C and D, used to evaluate the contribution from the lepton misidentification background.

There are 47 events selected in Region C and 606 selected in Region D, giving a scale factor of 0.078. The 105 events selected from Region B are scaled by this factor to give a total of 8.19 ± 2.32 expected lepton misidentification background events, where the uncertainty is taken as the statistical uncertainty of the number of events selected in each control region summed in quadrature.

5.4.2 Z + jet

Contributions from Z + jet events must be considered for cases in which the Z boson decays to two tau leptons which then produce a final state including a muon, electron and jet. A Feynman diagram of a Z + jet event is shown in Figure 5.8.

The contribution from $(Z \rightarrow \tau\tau)$ + jet events is evaluated by comparing $(Z \rightarrow \tau\tau)$ + jet and $(Z \rightarrow \mu\mu)$ + jet events in simulation to obtain a normalisation factor to determine the number of $(Z \rightarrow \tau\tau)$ + jet events expected for every $(Z \rightarrow \mu\mu)$ + jet event seen. The

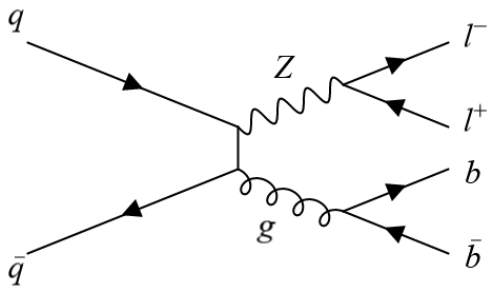

FIGURE 5.8: Feynman diagram of a Z + jet event.

factor is evaluated as 4.3×10^{-4}. The uncertainty on this value is negligible in comparison to the uncertainty arising from reconstruction and selection efficiencies described later, and so is not considered.

$(Z \to \mu\mu)$ + jet events are then selected in data using the selection criteria summarised in Table 5.5, and scaled using the normalisation factor. The selection requires at least one of the muons to have fired all three of the `L0EWMuon`, `HLT1SingleMuonHighPT` and `HLT2EWSingleMuonVHighPt` trigger lines. Both muons must have p_{T} greater than 20 GeV and be inside the LHCb acceptance. The invariant mass of the reconstructed Z boson must be between 60 and 120 GeV. The jet must have a p_{T} greater than 20 GeV, $2.2 < \eta < 4.2$ and be SV-tagged as described in Section 5.3.3.

A total of 2316 events are selected in data as shown in Figure 5.9, and scaled to give an expected background of 1.0 ± 0.1 events. The uncertainty assigned to this value is a conservative 10% to account for differences between data and MC in reconstruction and selection efficiencies as discussed in Section 5.5.

Trigger	`L0EWMuon`
	`HLT1SingleMuonHighPT`
	`HLT2EWSingleMuonVHighPt`

Muon	`isMuon` $p_{\mathrm{T}} > 20$ GeV, $2.0 < \eta < 4.5$
Boson	$60 < M_{\mu\mu} < 120$ GeV
Jet	$p_{\mathrm{T}} > 20$ GeV, $2.2 < \eta < 4.2$ SV-tagged

TABLE 5.5: A summary of the selection criteria applied to isolate $(Z \to \mu\mu)$ + jet candidates.

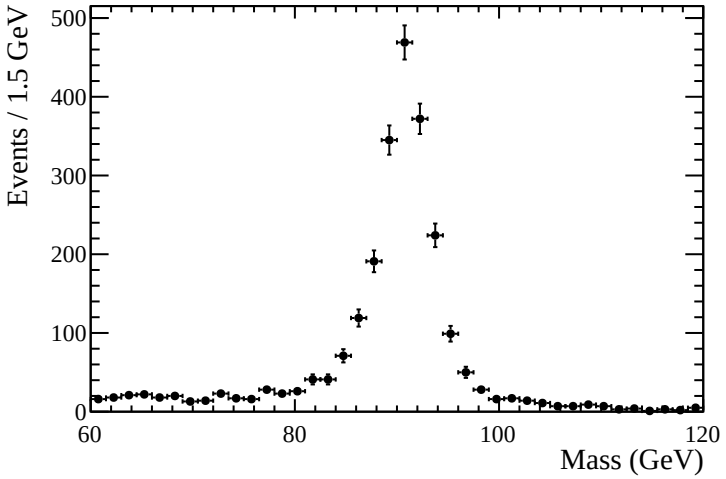

FIGURE 5.9: Invariant mass distribution of the $(Z \rightarrow \mu\mu)$ + jet events selected from the full Run-II dataset. These events are used to predict the Z + jet background contribution.

5.4.3 Single Top

The production of a single top quark in association with a W boson also produces a μeb final state and so can contribute a background with similar kinematics to $t\bar{t}$. A Feynman diagram displaying this interaction is shown in Figure 5.10.

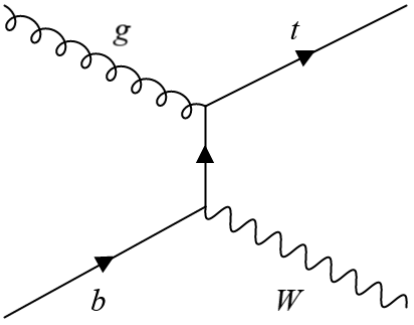

FIGURE 5.10: Feynman diagram of a single top event.

An estimation of the expected number of single top events in the signal sample is extrapolated from Reference [93], using simulation from POWHEG with both the *diagram removal* and *diagram subtraction* schemes outlined in Reference [97]. The $t\bar{t}$ MC sample is reweighted to determine the expected single top reconstruction and selection efficiencies as no full detector simulation of the process is available. From this method, a background of 5.09±1.53 events is expected, which includes the difference between the two schemes.

5.4.4 Di-boson

Pairs of W and Z bosons can decay to produce a final state with a muon and electron, which can contribute a background when a jet is present in the event. This contribution is modelled using aMC@NLO and the expected contribution from all final states is found to be below the per-mille level. This background is not considered further.

5.4.5 Misidentified Jets

Misidentified jets are incorrectly reconstructed jets, using fake reconstructed particles or particles from a different pp collision in the same event. These jets can contribute a background when two leptons are selected in association with a misidentified jet. These jets are not expected to contribute a significant background as they are typically softer than the p_{T} requirement. The contribution from such events is evaluated using simulation and determined to be negligible and so is not considered further.

5.4.6 Summary

A summary of the expected number of background events and the methods used to determine them is shown in Table 5.6.

Figures 5.11,5.12 and 5.13 show the selected data compared to estimated background contributions. The signal is displayed by MC normalised to the number of candidates selected in data minus the total number of expected background events. The single top contribution is modelled using $t\bar{t}$ MC normalised to the POWHEG prediction. The Z + jet shapes are taken from $Z \rightarrow \tau\tau$ MC for lepton distributions and $(Z \rightarrow \mu\mu)$ + jet MC for jet distributions.The lepton misidentification shapes are taken from the anti-electron identification control sample with the jet tagging requirement removed to improve statistical precision.

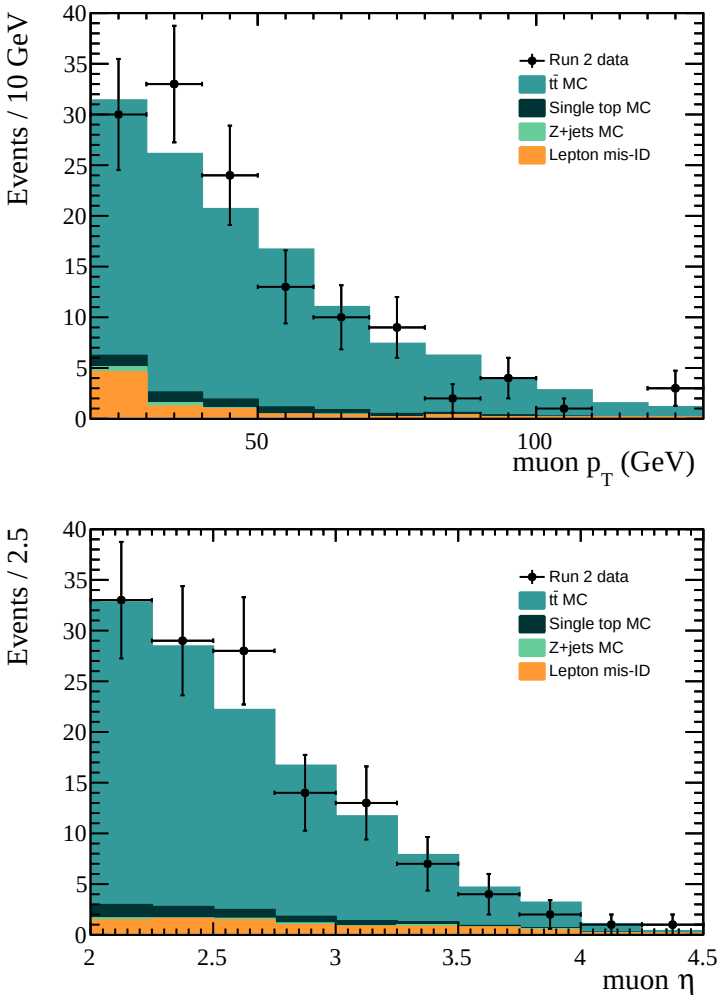

FIGURE 5.11: The muon p_{T} (top) and pseudorapidity (bottom) distributions in data compared to the expected background contributions.

Source	Method	Events
Lepton mis-identification	Same-sign data	8.2 ± 2.3
Z + jet	Normalised to $(Z \to \mu\mu)$ + jet	1.0 ±0.1
Single top	POWHEG predictions	5.1 ± 1.5
Di-boson	aMC@NLO predictions	negligible
Fake jets	PYTHIA simulation	negligible
Total		14.3±2.7

TABLE 5.6: A summary of the expected contributions from each background and the methods used to evaluate them.

5.5 Efficiencies

The efficiency with which events are selected, reconstructed and tagged is evaluated using the techniques outlined in the following section. Unless otherwise stated, the efficiencies

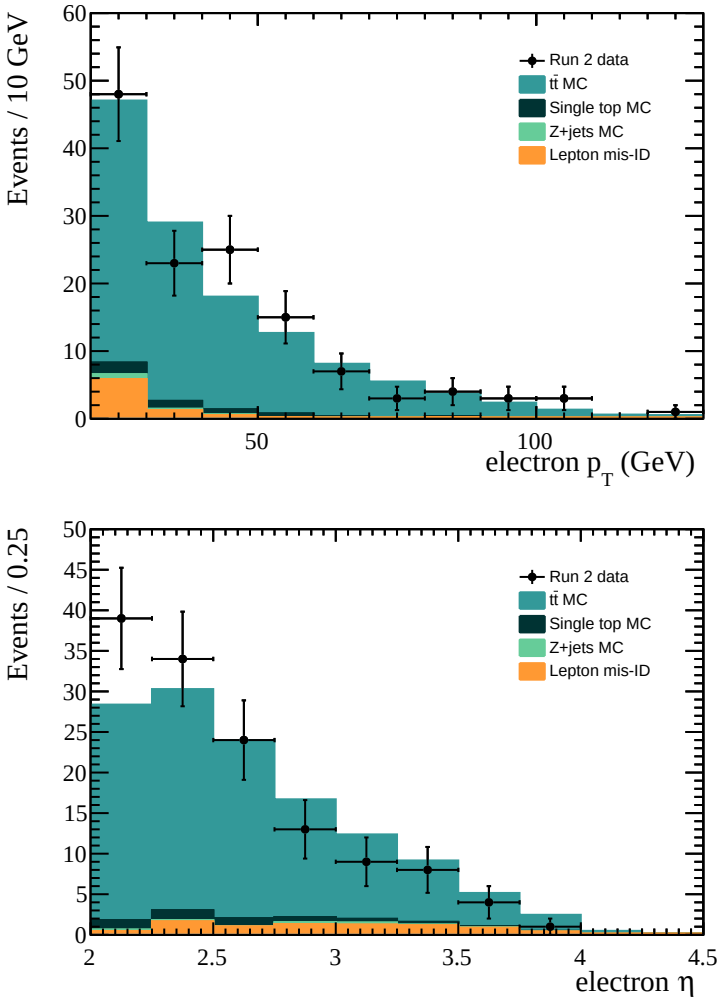

FIGURE 5.12: The electron p_{T} (top) and pseudorapidity (bottom) distributions in data compared to the expected background contributions.

are evaluated using simulation corrected to data.

5.5.1 Reconstruction Efficiencies

The reconstruction efficiency is defined as:

$$\varepsilon^{rec} = \varepsilon_{\mu}^{trg} \times \varepsilon_{\mu}^{id} \times \varepsilon_{\mu}^{trk} \times \varepsilon_{e}^{id} \times \varepsilon_{e}^{trk} \times \varepsilon_{jet}^{rec} \tag{5.3}$$

where ε_{μ}^{trg}, ε_{μ}^{id} and ε_{μ}^{trk} are the muon trigger, identification and tracking efficiencies; ε_{e}^{id} and ε_{e}^{trk} are the electron identification and tracking efficiencies; and ε_{jet}^{rec} is the jet reconstruction efficiency.

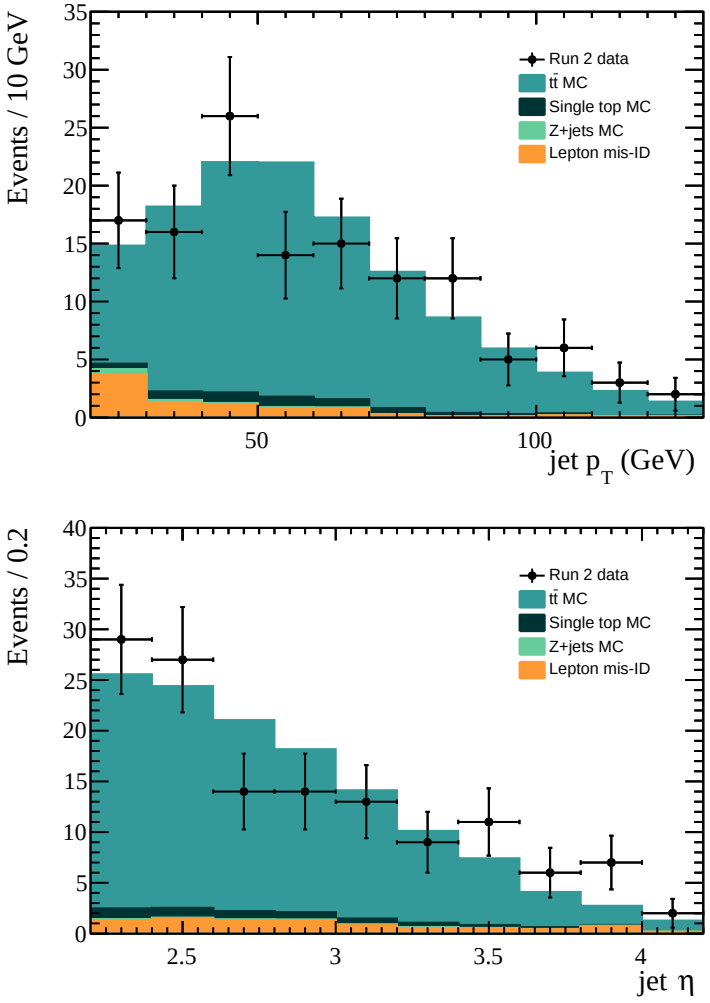

FIGURE 5.13: The jet p_{T} (top) and pseudorapidity (bottom) distributions in data compared to the expected background contributions.

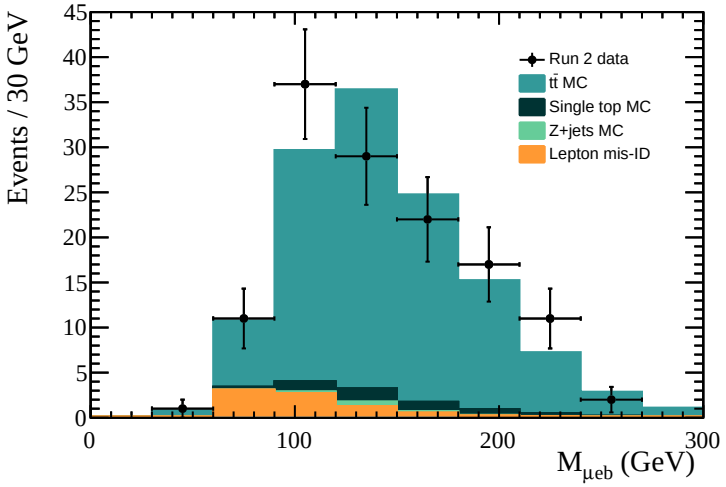

FIGURE 5.14: The invariant mass of the muon, electron and b-jet in data compared to the expected background contributions.

The lepton reconstruction efficiencies are evaluated using tag-and-probe methods in which data and MC samples of $Z \to \mu\mu$ and $Z \to ee$ are compared. In this method, one lepton is identified as a tag, having passed a given set of requirements. The number of tags passing the selection is taken as the total number of events. The second lepton, or probe, is then assessed against further requirements and the number of passes is counted. The efficiencies are calculated by dividing the number of passes by the total number of events selected, as shown in Equation 5.4:

$$\varepsilon = \frac{N_{pass}}{N_{total}} \tag{5.4}$$

where ε is the tag-and-probe efficiency, N_{pass} is the number of events passing the additional probe requirements and N_{total} is the total number of tags selected.

Details of the tag-and-probe methods used to calculate each of the lepton reconstruction efficiencies are presented in the following sections.

The technique used to determine the jet reconstruction efficiency is different and is described in detail in Section 5.5.1.6.

5.5.1.1 Trigger Efficiency

The trigger efficiency is calculated using $Z \to \mu\mu$ data where both muons have been identified using the `isMuon` criteria, similar to the method used in Reference [98]. The tag muon must pass each of the three trigger stages `L0EWMuon`, `HLT1SingleMuonHighPT` and `HLT2EWSingleMuonVHighPt` as described in Section 5.3.1. The probe muon is counted as a pass when it also passes each of the three trigger stages.

Further selection criteria are applied to ensure the purity of the sample. The leptons must both have p_{T} greater than 20 GeV, be within the pseudorapidity range $2 < \eta < 4.5$ and be isolated with a $p_{\mathrm{T}}^{\mathrm{Cone}}$ of less than 2 GeV. The di-muon invariant mass is required to be between 60 and 120 GeV and the χ^2_{vtx}/nDF of the primary vertex must be less than 5. This is summarised in Table 5.7.

The trigger efficiency is shown as a function of p_{T} and pseudorapidity for data and MC in Figure 5.15. The differences between the data and simulation are mostly due to the

efficiency of the hardware trigger, and the increased disparity at high η most likely being caused by alignment effects in the tracking system.

Tag	Probe	Event
Long track	Long track	$60 < M_{\mu\mu} < 120$ GeV
Triggered	-	$\chi^2_{vtx}/nDF < 5$
`isMuon`	`isMuon`	-
$p_{\mathrm{T}} > 20$ GeV	$p_{\mathrm{T}} > 20$ GeV	-
$p_{\mathrm{T}}^{\mathrm{Cone}} < 2$ GeV	$p_{\mathrm{T}}^{\mathrm{Cone}} < 2$ GeV	-
$2 < \eta < 4.5$	$2 < \eta < 4.5$	-

TABLE 5.7: A summary of the selection criteria applied to $Z \to \mu\mu$ data to calculate the trigger efficiency.

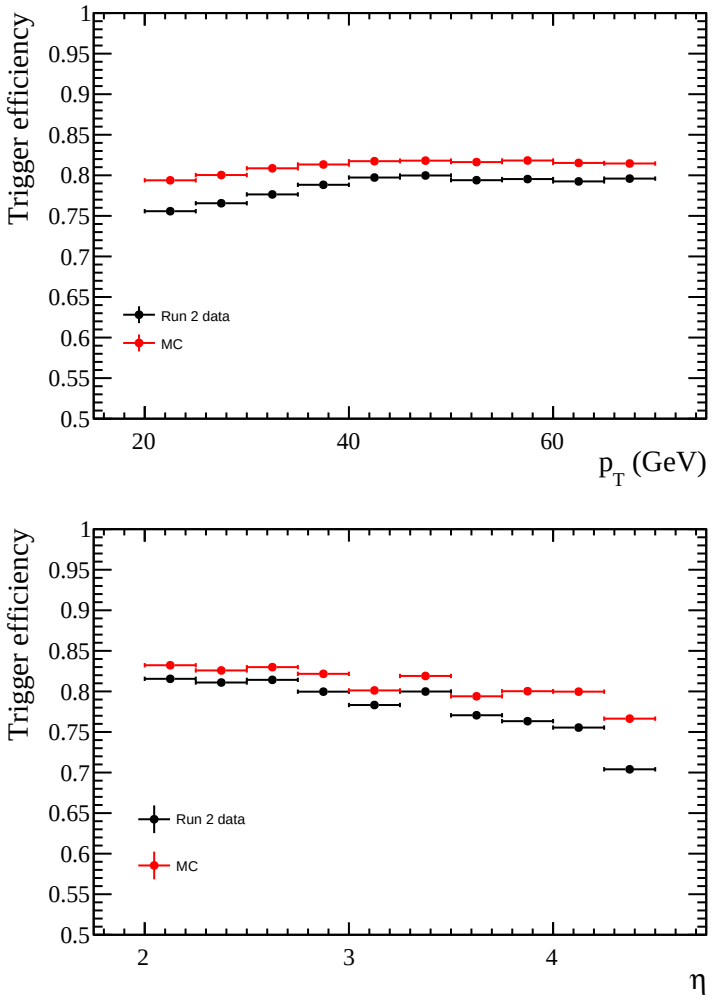

FIGURE 5.15: The muon trigger efficiency as a function of p_{T} (top) and η (bottom) for data and simulation.

The data and MC efficiencies are then re-calculated in two dimensions of η and p_{T} using 10×10 bins. The data efficiency is divided by the MC efficiency to calculate a

data/simulation correction for each (η, p_{T}) region.

A final value for the trigger efficiency is then calculated using existing methods, by applying the correction factors to $t\bar{t}$ simulation event by event and taking the overall pass rate as the efficiency. The trigger efficiency is found to be 0.813±0.010, where the uncertainty includes the difference in efficiency before and after the corrections are applied and the uncertainty due to available statistics.

5.5.1.2 Muon Identification Efficiency

The muon identification efficiency is also evaluated using $Z \rightarrow \mu\mu$ data and the method described in Reference [98]. This time, the tag is a muon identified using the `isMuon` criteria, which must also fire all three stages of the trigger. The probe is a long track as described in Section 4.1 and is counted as a pass when it satisfies the `isMuon` requirement.

The event selection criteria requires both the tag and probe to have p_{T} greater than 20 GeV, pseudorapidity in the range $2 < \eta < 4.5$, a $p_{\mathrm{T}}^{\mathrm{Cone}}$ of less than 2 GeV and to be separated by at least 2.7 radians in ϕ. The di-muon invariant mass is again required to be between 60 and 120 GeV and the χ^2_{vtx}/nDF of the primary vertex must be less than 5. This is summarised in Table 5.8.

A comparison of the muon identification efficiency in data and MC is shown as a function of p_{T} and pseudorapidity in Figure 5.16. The reduction in the data efficiency compared with simulation at low p_{T} is most likely due to increased background contribution at the p_{T} requirement threshold. The drop in efficiency in the highest pseudorapidity bin in both the data and MC distributions is due to the limited acceptance of the muon stations close to the beampipe.

Tag	Probe	Event
Long track	Long track	$60 < M_{\mu\mu} < 120$ GeV
Triggered	Triggered	$\chi^2_{vtx}/nDF < 5$
`isMuon`	-	$\lvert\Delta\phi\rvert > 2.7$ rad
$p_{\mathrm{T}} > 20$ GeV	$p_{\mathrm{T}} > 20$ GeV	-
$p_{\mathrm{T}}^{\mathrm{Cone}} < 2$ GeV	$p_{\mathrm{T}}^{\mathrm{Cone}} < 2$ GeV	-
$2 < \eta < 4.5$	$2 < \eta < 4.5$	-

TABLE 5.8: A summary of the selection criteria applied to $Z \rightarrow \mu\mu$ data to calculate the muon identification efficiency.

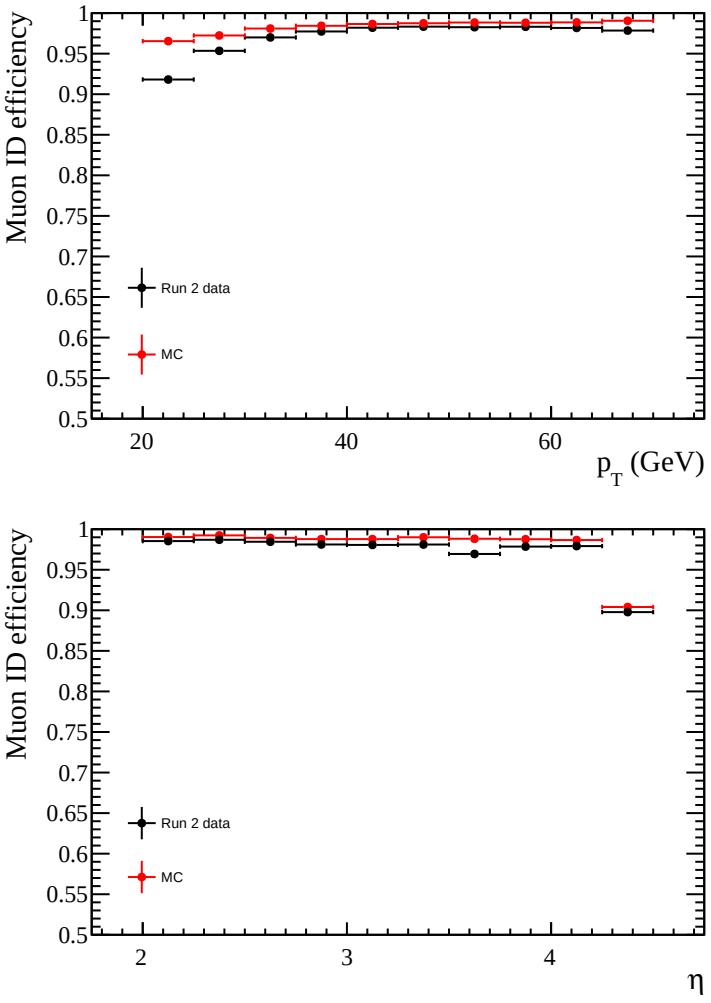

FIGURE 5.16: The muon identification efficiency as a function of p_T (top) and η (bottom) for data and simulation.

Data/simulation correction factors and a final muon identification efficiency value are calculated using a similar approach to that used to evaluate the the trigger efficiency. The muon identification is found to be 0.978±0.012.

5.5.1.3 Muon Tracking Efficiency

The muon tracking efficiency is also evaluated using $Z \rightarrow \mu\mu$ events and a tag-and-probe method built on the one described in Reference [98]. In this instance, the tag is a triggered and identified muon and the probe is a MuonTT track. A MuonTT track is a track reconstructed by combining hits in the muon stations with hits in the TT and is chosen as the probe because hits in the muon stations and TT are not included in long

track reconstruction. The probe is counted as a pass when a long track is reconstructed in association with the MuonTT track.

Events are selected by requiring the tag and probe muons to have p_{T} greater than 20 GeV, pseudorapidity in the range $2 < \eta < 4.5$, and to be separated by at least 0.1 radians in ϕ. The di-muon invariant mass is required to be between 70 and 110 GeV and the χ^2_{vtx}/nDF of the primary vertex must be less than 5. This selection is summarised in Table 5.9.

The muon tracking efficiency is shown for data and MC as a function of p_{T} and pseudorapidity in Figure 5.17. The differences between the data and simulation distributions are caused by material interactions and detector misalignment.

The data/simulation correction factors and a final value for the muon tracking efficiency are calculated using a similar method as for the trigger and muon identification efficiencies. The muon tracking efficiency is 0.935±0.020.

Tag	Probe	Event
Long track	MuonTT track	$70 < M_{\mu\mu} < 110$ GeV
Triggered	-	$\chi^2_{vtx}/nDF < 5$
`isMuon`	-	$\lvert\Delta\phi\rvert > 0.1$ rad
$p_{\mathrm{T}} > 20$ GeV	$p_{\mathrm{T}} > 20$ GeV	-
$2 < \eta < 4.5$	$2 < \eta < 4.5$	-

TABLE 5.9: A summary of the selection criteria applied to $Z \to \mu\mu$ data to calculate the muon tracking efficiency.

5.5.1.4 Electron Identification Efficiency

The electron identification efficiency is also calculated using a tag-and-probe method, this time using $Z \to ee$ data and MC. The technique used is based on the one used in Reference [95]. The tag is an identified and triggered electron, using the the electron identification criteria outlined in Section 5.3.2.3 and the `L0Electron`, `Hlt1SingleElectronNoIP` and `Hlt2EWSingleElectronVHighPt` trigger lines as defined in Section 4.4 and used in Reference [99]. The probe is a long track and is counted as a pass when it fulfills the same electron identification requirements.

The electrons are required to have a p_{T} of greater than 20 GeV, to be within the LHCb acceptance ($2 < \eta < 4.5$) and to be separated in ϕ by at least 2.7 radians. They

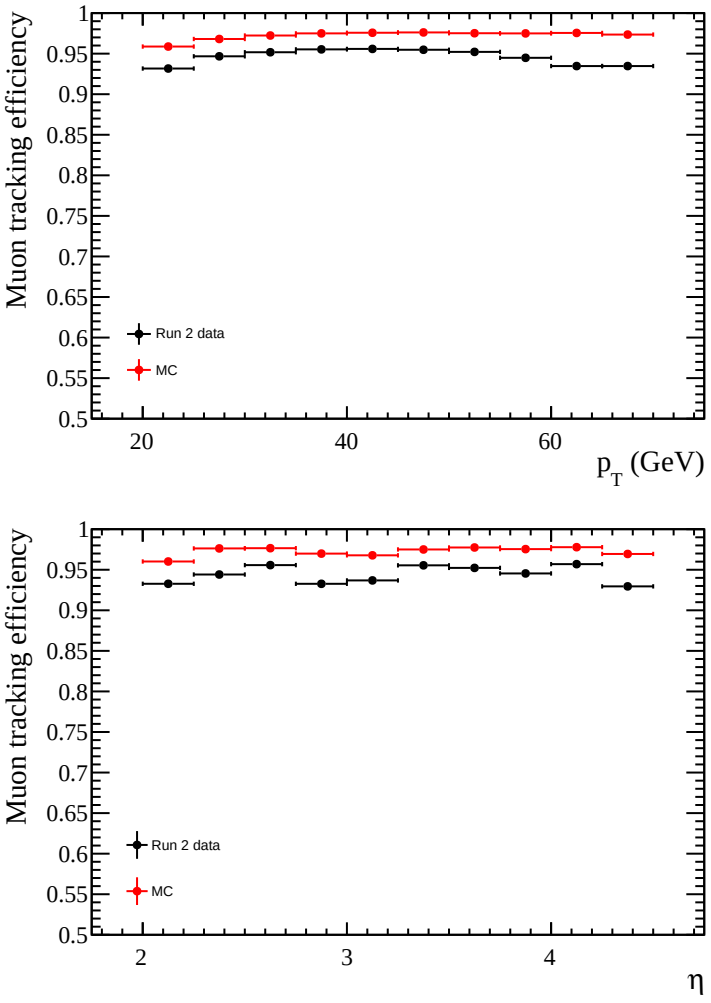

FIGURE 5.17: The muon tracking efficiency as a function of p_{T} (top) and η (bottom) for data and simulation.

must also be isolated using the $p_{\mathrm{T}}^{\mathrm{Cone}}$ variable defined in Equation 5.1 where the cone is annular with $0.1 < R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.5$. This is different to the muon selection due to photons emitted through Bremsstrahlung travelling close to the electron track. The invariant mass of the combined electrons is required to be between 70 and 110 GeV and the χ^2_{vtx}/nDF of the primary vertex must be less than 5. These requirements are summarised in Table 5.10.

A residual background of around 8% is observed after the selection is applied. This is taken into account by applying the same tag-and-probe procedure to a second $Z \rightarrow ee$ sample where the tag and probe are required to have the same charge. The same-sign contribution is then subtracted from both the total and pass histograms in the opposite-sign samples used to calculate the efficiency. An example of this background subtraction

can be seen in Figure 5.18.

The electron identification efficiency is compared in data and MC in Figure 5.19, where a single bin is used for events above 70 GeV in the p_{T} case. The drop in efficiency at high and low η is due to the acceptance of the ECAL.

Data/simulation corrections and a final value for the electron identification efficiency are calculated as for each of the muon reconstruction efficiencies. The electron identification efficiency is found to be 0.913±0.044.

Tag	Probe	Event
Long track	Long track	$70 < M_{\mu\mu} < 110$ GeV
Triggered	-	$\chi^2_{vtx}/nDF < 5$
`!isMuonLoose`	-	$\lvert\Delta\phi\rvert > 2.7$
$p_{\mathrm{T}} > 20$ GeV	$p_{\mathrm{T}} > 20$ GeV	-
$2 < \eta < 4.5$	$2 < \eta < 4.5$	-
$p_{\mathrm{T}}^{\mathrm{Cone}} < 2$ GeV	$p_{\mathrm{T}}^{\mathrm{Cone}} < 2$ GeV	-
$E_{PRS} > 50$ MeV	-	-
$E_{HCAL}/P < 0.05$	-	-
$E_{ECAL}/P > 0.1$	-	-

TABLE 5.10: A summary of the selection criteria applied to $Z \rightarrow ee$ data to calculate the electron identification efficiency.

5.5.1.5 Electron Tracking Efficiency

The efficiency with which electron tracks are reconstructed is evaluated using the same method as Reference [93]. Unlike in the muon case, an unbiased probe cannot be constructed for electron tracking. Instead, two tag-and-probe samples are used to determine the efficiency: a $Z \rightarrow ee$ sample where the tag is a triggered and identified electron and the probe is a long track; and a $Z \rightarrow e\gamma$ sample where the tag is a triggered and identified electron and the probe is a photon. The $Z \rightarrow e\gamma$ sample acts as an estimate for how often electron tracks are "missed" in the reconstruction, i.e. when an electron is present, but a track has not been reconstructed by the tracking algorithms.

In the $Z \rightarrow e\gamma$ sample, the tag must pass the `L0Electron`, `Hlt1SingleElectronNoIP` and `Hlt2EWSingleElectronVHighPt` trigger lines. It must also satisfy the electron identification requirements outlined in Section 5.3.2.3 and have p_{T} greater than 20 GeV, $2 < \eta < 4.5$ and $p_{\mathrm{T}}^{\mathrm{Cone}}$ less than 2 GeV. The probe is a reconstructed photon with E_{T} greater than 20 GeV, $2 < \eta < 4.0$ and $p_{\mathrm{T}}^{\mathrm{Cone}}$ less than 2 GeV. Here the cone is annular

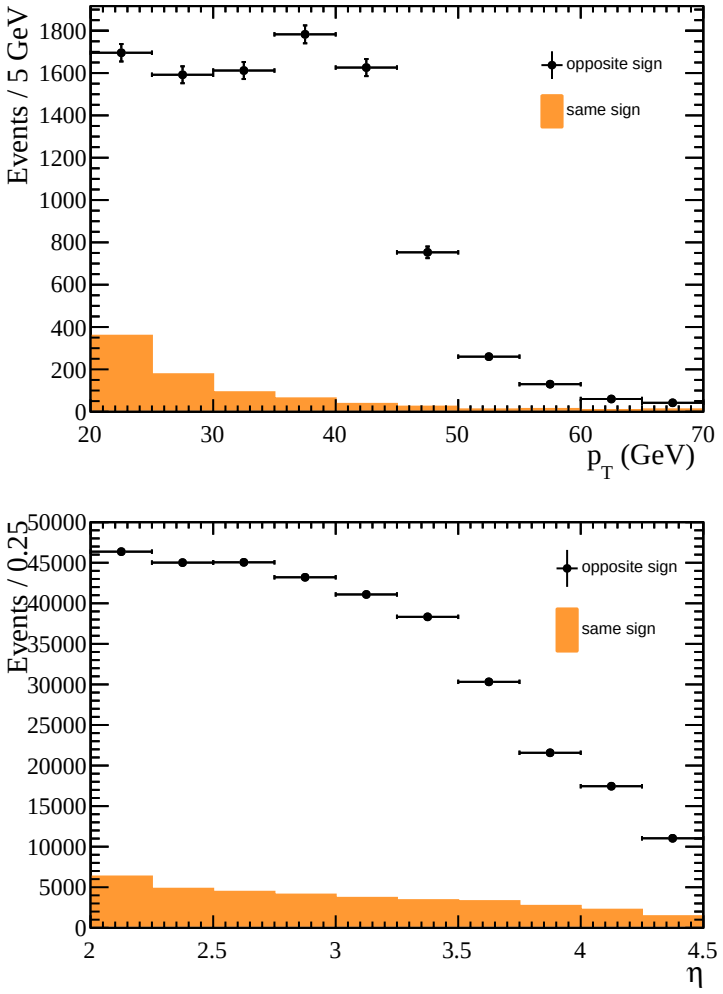

FIGURE 5.18: Plots to represent the same-sign data subtraction from opposite-sign events while evaluating the electron identification efficiency in electron p_T (top) and pseudorapidity (bottom) distributions.

with radius $R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ between 0.1 and 0.5 as for the electron identification efficiency selection criteria outlined in Section 5.5.1.4. The tag and probe must be separated by 2.7 radians in ϕ. The selection is the same for the $Z \rightarrow ee$ sample, but with the additional requirement of an ECAL cluster with $E_T > 10$ GeV associated with the probe track. This ensures the calorimeter cluster requirements are the same as for the photon probe sample. The selection requirements are summarised in Tables 5.11 and 5.12. As for the electron identification efficiency, an additional background is considered for the $Z \rightarrow ee$ sample by considering a same-sign data sample and removing the same number of events from the opposite-sign data.

A significant background is admitted through the tag-and-probe selection so the number of signal events is determined from an existing template fit method. A background

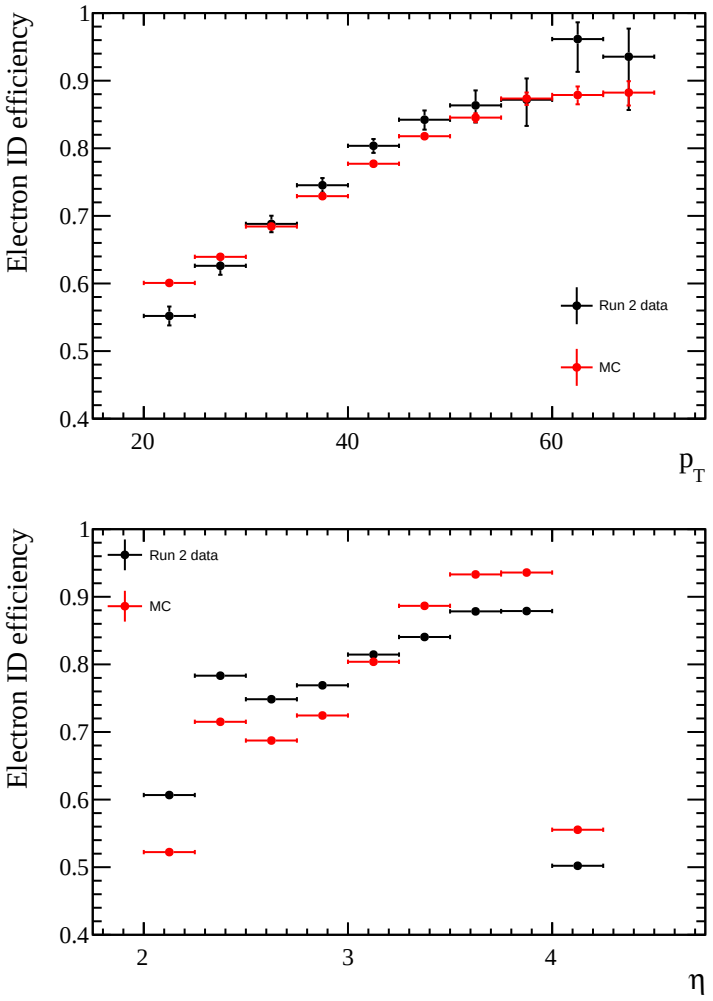

FIGURE 5.19: The electron identification efficiency as a function of p_T (top) and η (bottom) for data and simulation.

Tag	Probe	Event
Long track	Long track	$\lvert\Delta\phi\rvert > 2.7$
Triggered	-	-
`!isMuonLoose`	-	-
$p_T > 20$ GeV	$E_T > 20$ GeV	-
$2 < \eta < 4.5$	$2 < \eta < 4.0$	-
$p_T^{Cone} < 2$ GeV	$p_T^{Cone} < 2$ GeV	-
$E_{PRS} > 50$ MeV	-	-
$E_{HCAL}/P < 0.05$	-	-
$E_{ECAL}/P > 0.1$	-	-

TABLE 5.11: A summary of the selection criteria applied to $Z \to e\gamma$ events to calculate the electron tracking efficiency.

shape is taken from each sample with the same tag-and-probe selection, except the isolation requirement is reversed with p_T^{Cone} required to be greater than 5 GeV. The signal

Tag	Probe	Event
Long track	Long track	$\lvert\Delta\phi\rvert > 2.7$
Triggered	-	-
`!isMuonLoose`	-	-
$p_{\mathrm{T}} > 20$ GeV	$p_{\mathrm{T}} > 20$ GeV	-
$2 < \eta < 4.5$	$2 < \eta < 4.0$	-
$p_{\mathrm{T}}^{\mathrm{Cone}} < 2$ GeV	$p_{\mathrm{T}}^{\mathrm{Cone}} < 2$ GeV	-
$E_{PRS} > 50$ MeV	-	-
$E_{HCAL}/P < 0.05$	-	-
$E_{ECAL}/P > 0.1$	$E_{ECAL} > 10$ GeV	-

TABLE 5.12: A summary of the selection criteria applied to $Z \rightarrow ee$ events to calculate the electron tracking efficiency.

component in each sample is then determined using the shape from the $Z \rightarrow ee$ sample, where the probe is additionally required to pass the electron identification criteria. The tag p_{T} distributions for both samples can be seen in Figure 5.20. The slight shift observed between the data and model in the $Z \rightarrow e\gamma$ plot is due to the imperfect description of the background and is covered by the systematic uncertainty of the fit.

The same process is then repeated with simulated samples of $Z \rightarrow e\gamma$ and $Z \rightarrow ee$ as seen in Figure 5.21.

The electron tracking efficiency is estimated by dividing the total number of events selected in the $Z \rightarrow ee$ sample by the sum of the signal events found in the $Z \rightarrow ee$ and $Z \rightarrow e\gamma$ samples. This calculation is performed separately for the data and MC samples, with the efficiency in data found to be 0.899 ± 0.025 and the efficiency in MC evaluated as 0.923 ± 0.029, where the uncertainties are taken as the statistical uncertainty[1] and the uncertainty from the fitting procedure added in quadrature with the difference between the two central values. These values are then compared with the truth level[2] efficiency of 0.914. The electron tracking efficiency is taken from truth, with the uncertainty from the data/simulation metrics added in quadrature. It is found to be 0.914 ± 0.028.

[1] The systematic uncertainty is large enough to cover the small uncertainty associated with the same-sign background subtraction.

[2] Truth level simulations are simulations before any detector interaction or performance is included. In contrast, simulated events that include the detector response are referred to as *reconstructed level* simulation where necessary.

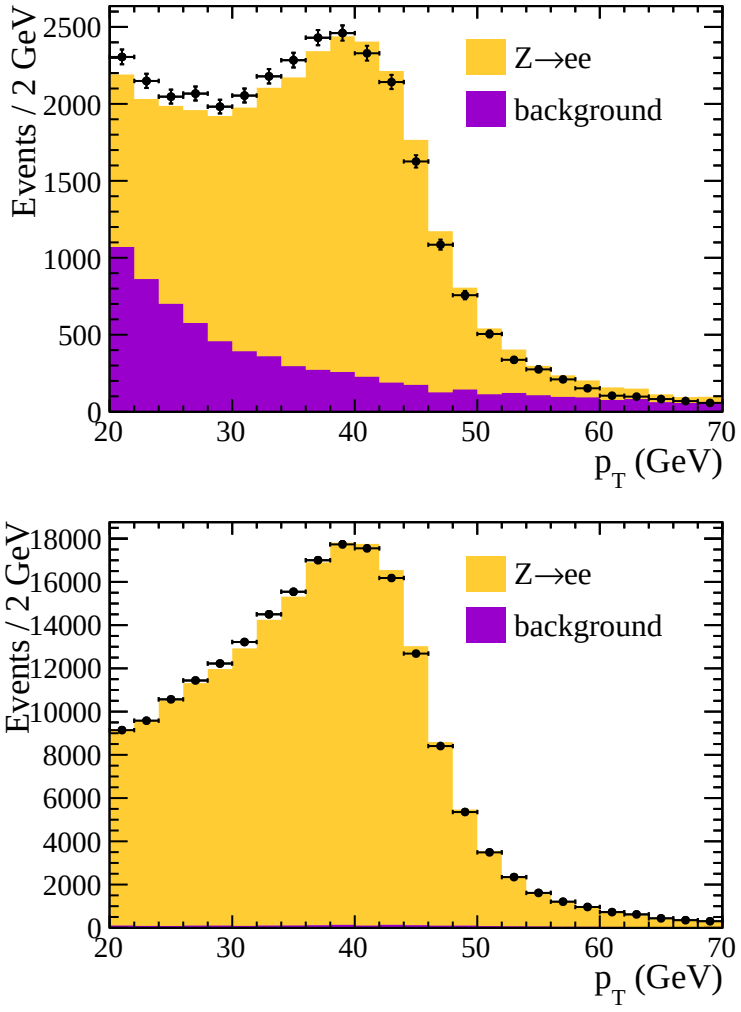

FIGURE 5.20: Template fits to $Z \to e\gamma$ (top) and $Z \to ee$ (bottom) data used in evaluating the electron tracking efficiency.

5.5.1.6 Jet Reconstruction Efficiency

The jet reconstruction efficiency is defined as the efficiency of reconstructing and identifying an MC truth level jet in the fiducial region as discussed in Reference [100]. The value is taken from simulation as the proportion of events passing this criteria in $t\bar{t}$ MC.

The uncertainty on this value is calculated by comparing the ratio of events when the nominal jet identification requirements, as defined in Reference [100], are applied and when individual requirements are tightened in data and simulation. This investigation is performed using $(Z \to \mu\mu)$ + jet events where the muons must be inside the LHCb acceptance and have a p_{T} greater than 20 GeV; the Z boson mass must be between 60 and 120 GeV; the jet must have a p_{T} greater than 20 GeV and pseudorapidity between

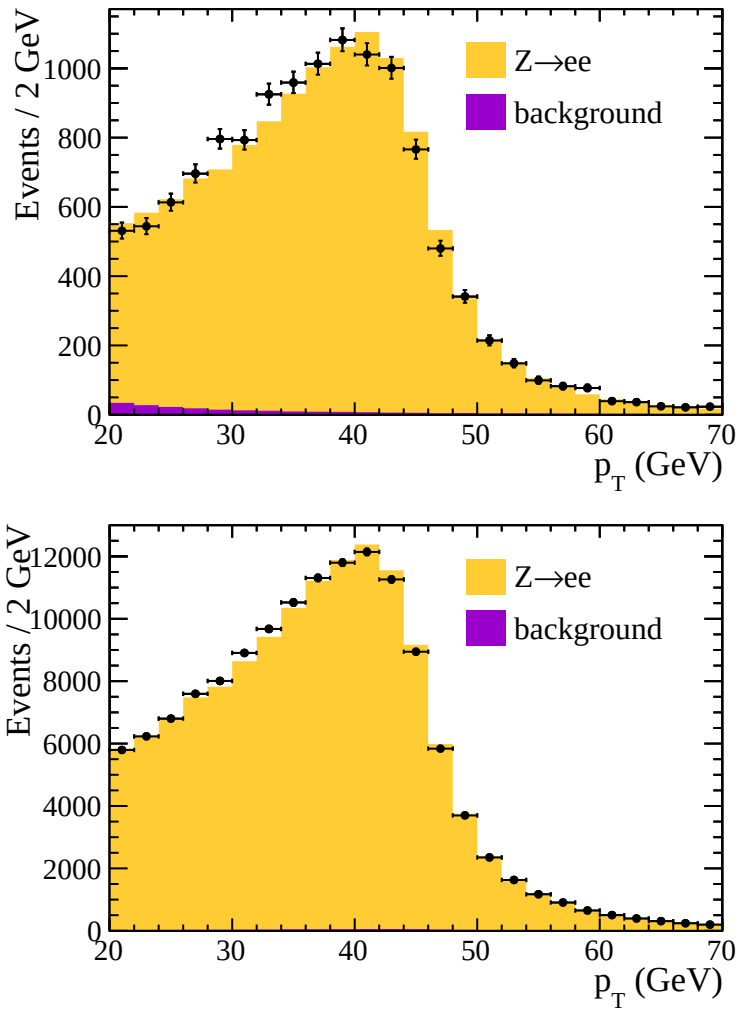

FIGURE 5.21: Template fits $Z \to e\gamma$ (top) and $Z \to ee$ (bottom) MC used in evaluating the electron tracking efficiency.

2.2 and 4.2; and at least one muon is required to pass all three stages of the trigger. This selection is summarised in Table 5.13.

Trigger	`L0EWMuon`
	`HLT1SingleMuonHighPT`
	`HLT2EWSingleMuonVHighPt`

Muon	`isMuon` $p_{\mathrm{T}} > 20$ GeV, $2.0 < \eta < 4.5$
Boson	$60 < M_{\mu\mu} < 120$ GeV
Jet	$p_{\mathrm{T}} > 20$ GeV, $2.2 < \eta < 4.2$

TABLE 5.13: A summary of the selection criteria applied to isolate $(Z \to \mu\mu)$ + jet candidates to evaluate the uncertainty on the jet reconstruction efficiency.

Four jet requirements are tightened in turn and the ratio of events passing the new

requirement for each change is compared between data and MC. The changes investigated are:

- increasing the minimum fraction of the jet composed of charged particles, cpf, from 10 to 20%,
- increasing the minimum p_{T} of the leading track in the jet, mpt, from 1.2 to 2.4 GeV,
- reducing the maximum fraction of the jet energy carried by a single particle, mtf, is reduced from 80 to 60%,
- increasing the minimum number of tracks which point to the primary vertex of the jet, $\mathrm{n_{pointing}}$, from 2 to 3.

Distributions of each variable before the requirements are tightened can be compared for data and MC in Figures 5.22 and 5.23. The difference in the data/simulation ratios for the efficiency of each change is added in quadrature to give the uncertainty. The jet reconstruction efficiency is 0.975 ± 0.013.

5.5.2 Jet Tagging Efficiency

To select events with b-jets for this analysis, a tagger is used to require that the jets in selected events are produced in a secondary vertex, as described in Section 5.3.3. A detailed comparison of the performance of this tagger in data and simulation was performed in Run-I, as outlined in Reference [96]. This comparison involved selecting events enriched in b and c jets by either requiring that the jets contain a muon or are observed in addition to a fully reconstructed B or D meson. To determine the efficiency of the jet tagging, a combined fit to all samples was used to estimate the flavour content of the jets before and after applying the SV-tagger. A conservative uncertainty of 10% was applied to account for differences between data and simulation, mostly observed at low p_{T}.

The jet tagging efficiency was assessed again in Run-II as discussed in [93]. The efficiency in the 2016 MC samples showed a relative reduction of 10% with respect to the 2012 MC. The jet tagging efficiency was validated using samples of $B \to J/\psi K$ events where the J/ψ was reconstructed from its decay to a pair of muons. The number of events

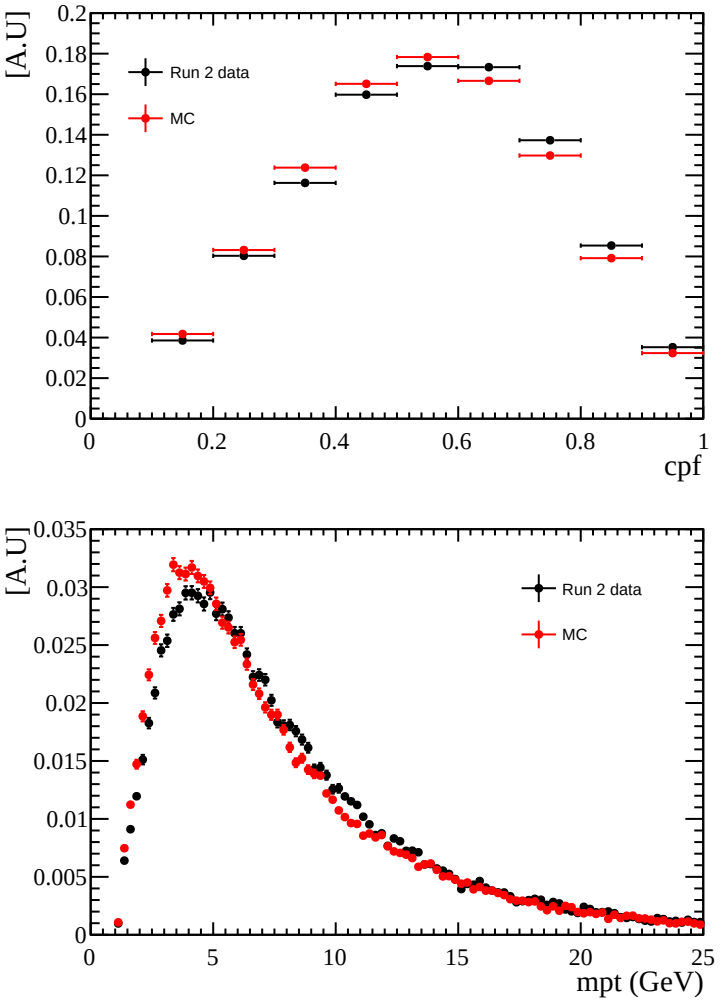

FIGURE 5.22: Distributions of jet quality variables used in assessing the uncertainty on the jet reconstruction efficiency: the charged particle fraction of the jet (top) and p_{T} of the main particle in the jet (bottom).

selected before and after the tagging requirements were applied was determined for each sample from an unbinned maximum likelihood fit to the J/ψ mass spectrum in different bins of phase space. Overall, a similar agreement between data and MC was seen for Run-I and Run-II and so the jet tagging efficiency was taken from 2016 MC, expected to be accurate to the same level as reported in Run-I.

As this analysis uses the same MC generation as the 2016 result, the jet tagging efficiency is again taken from 2016 simulation, with the same conservative uncertainty of 10% applied. This is validated by comparing the modelling of key variables used in the jet tagging process in data and simulation.

Events are selected in $(W \rightarrow \mu\nu)$ + jet data and MC where the muon is required to pass

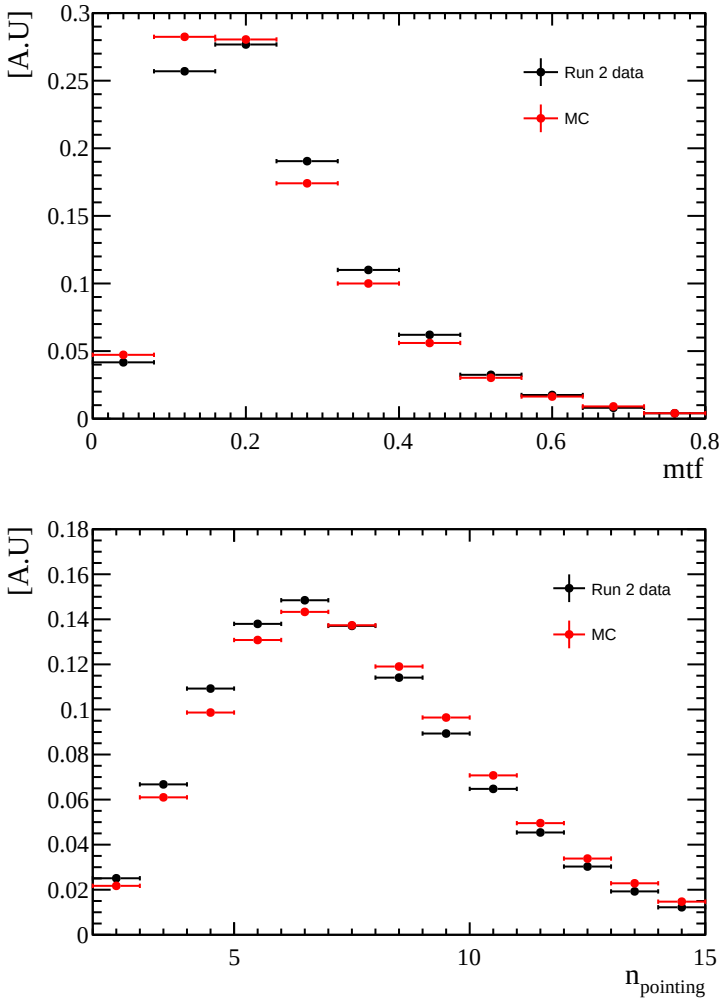

FIGURE 5.23: Distributions of jet quality variables used in assessing the uncertainty on the jet reconstruction efficiency: the fraction of the jet energy carried by the main particle in the jet (top) and the number of tracks with information pointing to the primary vertex (bottom).

all three stages of the trigger, have p_{T} greater than 20 GeV, ip less than 0.04 mm and be within the LHCb acceptance. The jet must be SV-tagged with p_{T} greater than 20 GeV, $2.2 < \eta < 4.2$, and must be separated from the muon by a radius of 0.5 in $\eta - \phi$ space. An isolation requirement is also placed on the muon, as described in Reference [100]. This is defined by replacing the muon with a jet, μ-jet, which contains the signal muon after a jet reconstruction has been performed. The p_{T} of the muon is required to account for at least 90% of the p_{T} of the μ-jet. The full selection is summarised in 5.14.

The key tagging variables are tightened in turn and the efficiency of each cut is evaluated for MC and for each year of data taking. The following variables are investigated:

Trigger	`L0EWMuon`
	`HLT1SingleMuonHighPT`
	`HLT2EWSingleMuonVHighPt`

Muon	`isMuon` $p_{\mathrm{T}} > 20$ GeV, $2.0 < \eta < 4.5$ $ip < 0.04$ mm, $p_{\mathrm{T}}/p_{\mathrm{T}}^{\mu-jet} > 0.9$
Jet	$p_{\mathrm{T}} > 20$ GeV, $2.2 < \eta < 4.2$, $\Delta R(\mu, j) > 0.5$ SV-tagged

TABLE 5.14: A summary of the selection criteria applied to $(W \to \mu\nu)$ + jet data to investigate the jet tagging efficiency.

- The limit on the minimum radial flight distance of the SV with respect to the PV, $\mathrm{FD}_{\mathrm{min}}$, is reduced from 15 to 10 mm,
- The minimum χ^2 of the flight distance for the SV, χ^2_{FD}, is tightened from 35 to 32,
- The lifetime of the SV, τ, is tightened from less than 1.5 ps to less than 1.2 ps,
- The maximum z-coordinate of the SV is reduced from 200 to 170 mm.

Differences in efficiency between each year and the simulation for each tagging variable are observed at the per mille level. A comparison of the variable distributions in data and MC can be seen in Figures 5.24 and 5.25. The differences seen between the data and simulation distributions are attributed to the simulation containing a higher light flavour contribution than would be expected in the data. In particular, this can be seen in the $\mathrm{FD}_{\mathrm{min}}$ distribution, where the MC peak at $\sim$5 mm is thought to be due to light flavour jet interaction with the VELO RF foil.

The efficiency of applying the SV tagger is also considered by comparing the ratio of events selected with and without the tagger applied for each year in data. Differences between the samples are at the per mille level.

The jet tagging efficiency is taken from 2016 MC. As the differences between this sample and the data are comparatively small, the same 10% uncertainty is applied as for previous analyses. The jet tagging efficiency is determined to be 0.556 ± 0.056.

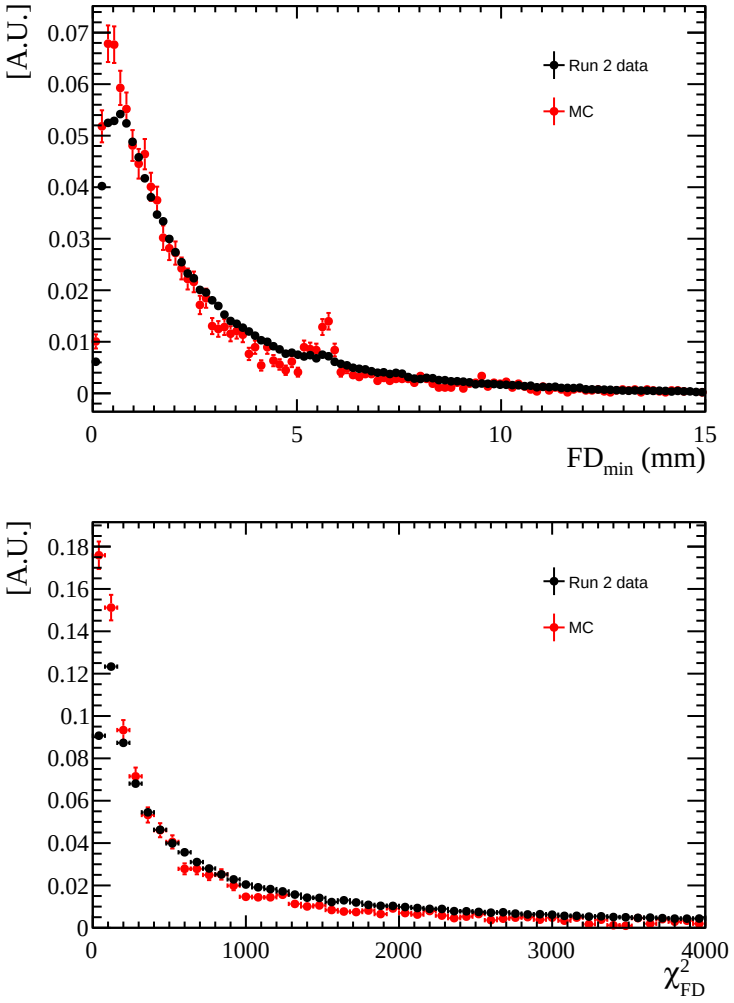

FIGURE 5.24: Tagging variables investigated for jet tagging efficiency: the minimum flight distance (top) and flight distance χ^2 (bottom).

5.5.3 Selection Efficiency

The selection efficiency refers to the efficiency of the isolation and impact parameter requirements used in the selection to isolate the $t\bar{t}$ signal.

The efficiency of the isolation requirement is taken directly from $t\bar{t}$ simuation as the ratio of events remaining in the sample after the $p_{\mathrm{T}}^{\mathrm{Cone}}$ cut is applied. Two sources of uncertainty on this value are considered: the modelling of the $t\bar{t}$ process in simulation and the level of agreement between simulation and data. To assess the former, the efficiency is calculated in different subsamples where the $t\bar{t}$ production is due to $q\bar{q}$ annihilation or gluon-gluon fusion, and where the final state contains one or multiple jets. The maximum difference in efficiency between the central value and the four samples is applied as a systematic uncertainty. The agreement between data and simulation is considered by

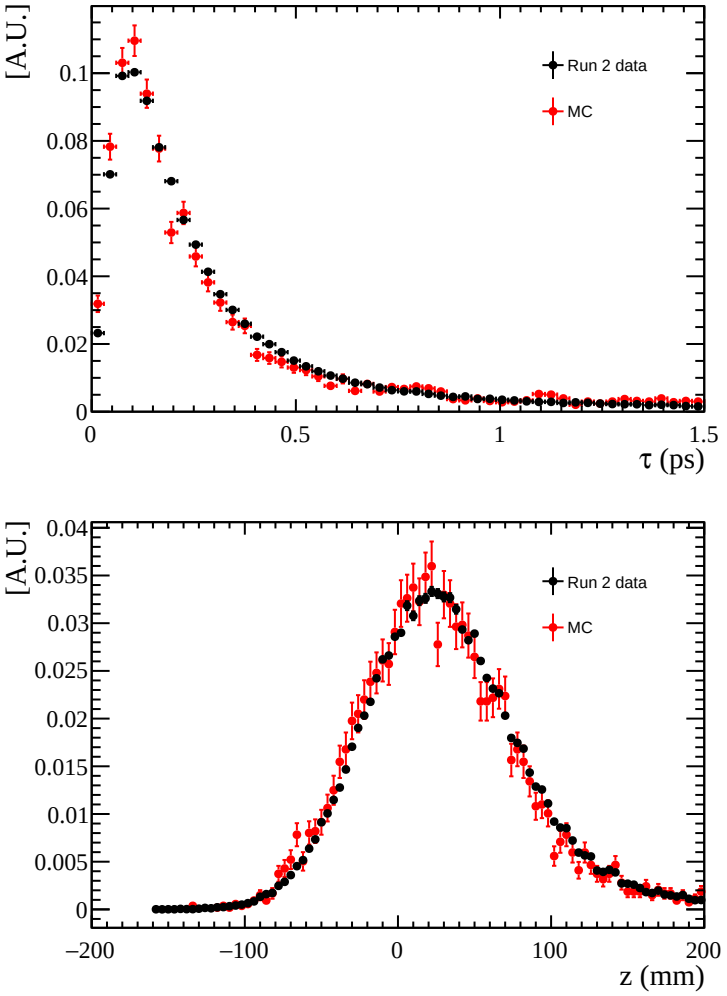

FIGURE 5.25: Tagging variables investigated for jet tagging efficiency: SV lifetime (top) and z position (bottom).

comparing the efficiency in $Z \rightarrow \mu\mu$ events, with the difference between the two values taken as an additional systematic.

As an imperfect agreement is observed in the modelling of the impact parameter, a tuning is performed to improve the description of the simulation before evaluating the efficiency of the *ip* requirement.

The impact parameter tuning requires using $Z \rightarrow \mu\mu$ data events in data and simulation and using the differences in the *ip* distributions to smear and shift the simulation to better model the data. The x and y components of the impact parameter are smeared by a Gaussian in 5×6 bins of η and ϕ. The mean and width of the Gaussian in each bin are determined by comparing simulation to data and performing a χ^2 minimisation using existing methods.

To assess the success of the tuning, *ip* distributions in data, MC and tuned MC are compared for the $Z \to \mu\mu$ events. The events in each sample are selected by requiring at least one of the muons to pass all three stages of the trigger and both muons to have a p_{T} greater than 20 GeV and to be inside the LHCb acceptance. The reconstructed Z boson is also required to have an invariant mass between 60 and 120 GeV. The selection criteria is summarised in Table 5.15 and the improvement to the impact parameter distribution after the tuning is applied can be seen in Figure 5.26.

Trigger	`L0EWMuon`
	`HLT1SingleMuonHighPT`
	`HLT2EWSingleMuonVHighPt`

Muon	`isMuon` $p_{\mathrm{T}} > 20$ GeV, $2.0 < \eta < 4.5$
Boson	$60 < M_{\mu\mu} < 120$ GeV

TABLE 5.15: A summary of the selection criteria applied to isolate $Z \to \mu\mu$ data and simulation to validate the impact parameter tuning required to evaluate the selection efficiency.

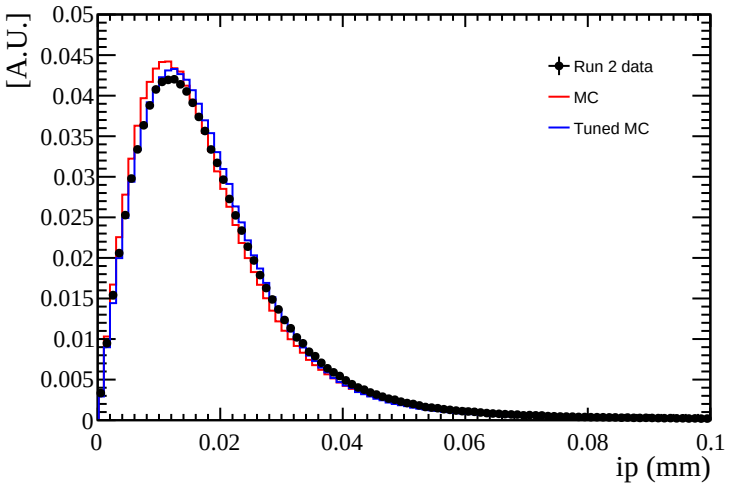

FIGURE 5.26: The impact parameter distribution of the muons in $Z \to \mu\mu$ events compared between data, MC and tuned MC.

The same tuning is then applied to the leptons in the $t\bar{t}$ MC event by event. The tuned $t\bar{t}$ MC is used to find the efficiency of the impact parameter cut. The efficiency of the *ip* requirement is taken as the central value with a systematic uncertainty applied as half of the difference between the efficiency before and after the tuning. Distributions of the muon and electron *ip* before and after the tuning can be seen in Figure 5.27.

The selection efficiency is evaluated as 0.666 ± 0.020.

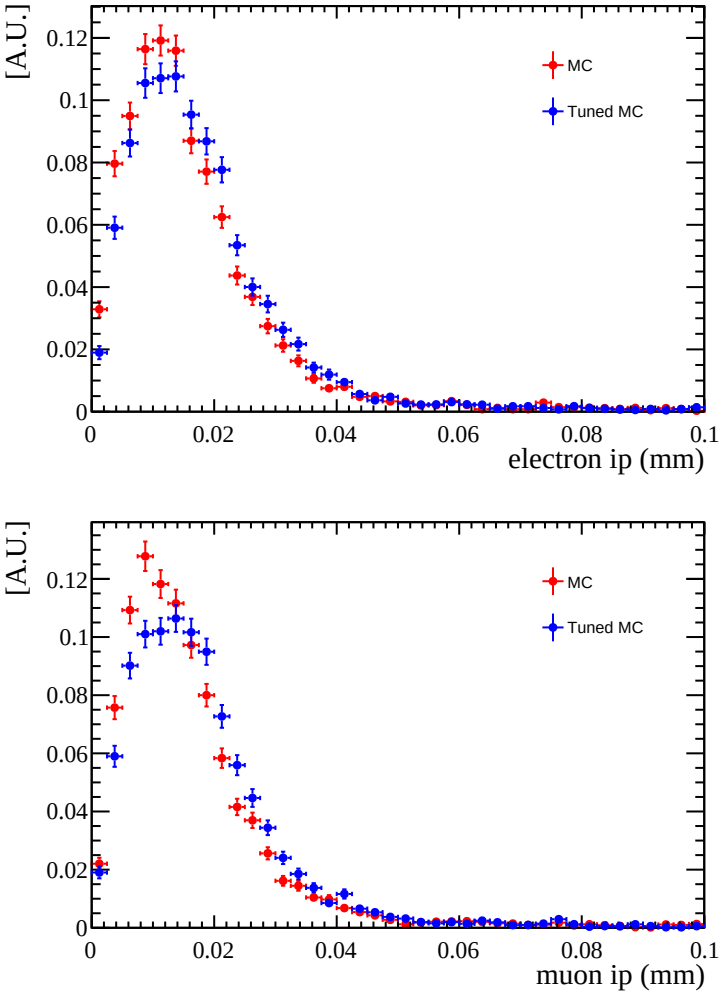

FIGURE 5.27: The impact parameter distributions before and after tuning for electrons (top) and muons (bottom) in $t\bar{t}$ MC.

5.5.4 Summary

The reconstruction, tagging and selection efficiencies described in the text are summarised in Table 5.16.

5.6 Resolution Factor

The resolution factor accounts for migrations into and out of the fiducial region due to the finite resolution of the detector. It is mainly affected by the p_{T} resolution of the reconstructed jet and the smearing of electron p_{T} to lower values due to incomplete Bremsstrahlung recovery.

Trigger	0.813±0.010
Muon identification	0.978±0.012
Muon tracking	0.935±0.020
Electron identification	0.913±0.044
Electron tracking	0.914±0.028
Jet reconstruction	0.975±0.013
Jet tagging	0.566±0.056
Selection	0.666±0.020
Total	0.224±0.028

TABLE 5.16: A summary of the efficiency to reconstruct, select and tag $t\bar{t} \to \mu eb$ events.

The resolution factor is calculated from simulation by comparing $t\bar{t}$ MC yields for events where the final state particles pass the reconstruction, identification and trigger requirements at truth and reconstructed levels. This is shown by Equation 5.5,

$$\mathcal{F}_{res} = \frac{N_{true}}{N_{rec}}, \tag{5.5}$$

where $\mathcal{F}_{res}$ is the resolution factor, N_{true} is the number of events expected at truth level and N_{rec} is the number of reconstructed MC events. The calculation of the resolution factor also includes a two-dimensional NLO re-weighting in electron and jet p_{T}, as they are the most significant variables.

The systematic uncertainty on the factor is considered by re-evaluating the value without the NLO reweighting and when corrections are applied to the jet and electron p_{T} scales to allow for differences in resolution between data and simulation. The differences in the resolution factor with and without corrections are considered as uncertainties as described in the following section.

5.6.1 Jet Energy Scale

The difference between the jet energy scale in data and MC contributes to the uncertainty on the resolution factor. This agreement is investigated using $(Z \to \mu\mu)$ + jet data and MC. The Z boson and jet in each event are produced approximately back-to-back in the transverse plane and so their p_{T} should be balanced. Therefore, the difference in p_{T} between the jet and Z boson can be used as a measure of the jet resolution.

To quantify this, a variable known as the jet 'balance' is defined as,

$$B = \frac{p_{\mathrm{T}Z} - p_{\mathrm{T}\mathrm{jet}}}{p_{\mathrm{T}Z}}, \tag{5.6}$$

where B is the balance, p_{T_Z} is the Z boson p_{T} and $p_{\mathrm{T}_{\mathrm{jet}}}$ is the p_{T} of the jet.

Both muons in the event are required to have p_{T} greater than 20 GeV, $2.0 < \eta < 4.5$, $p_{\mathrm{T}}^{\mathrm{Cone}}$ less than 5 GeV and ip less than 0.04 mm. At least one of the muons must pass all three stages of the trigger. The jet must have p_{T} greater than 20 GeV and $2.2 < \eta < 4.2$. Additionally, the jet is required to be separated from the Z boson by at least 2.7 radians in ϕ and from each lepton by a radius of 0.5 in $\eta - \phi$ space. This selection is outlined in Table 5.17 and a comparison of the distributions is shown in Figure 5.28. The differences between the data and MC distributions is then evaluated as a contibution to the uncertainty of the resolution factor.

Trigger	`L0EWMuon`
	`HLT1SingleMuonHighPT`
	`HLT2EWSingleMuonVHighPt`

Muons	`isMuon` $p_{\mathrm{T}} > 20$ GeV, $2.0 < \eta < 4.5$ $ip < 0.04$ mm, $p_{\mathrm{T}}^{\mathrm{Cone}} < 5$ GeV $\Delta R(\eta, e) > 0.1$
Jet	$p_{\mathrm{T}} > 20$ GeV, $2.2 < \eta < 4.2$, $\Delta R(l, j) > 0.5$, $\lvert\Delta\phi\rvert(Z, j) > 2.7$

TABLE 5.17: A summary of the selection criteria applied to $(Z \to \mu\mu)$ + jet data to calculate the jet balance.

The same method is used to plot the balance in ten bins of jet p_{T} in data and MC. A Crystal Ball fit is used to model the shape of the balance in each bin. A Crystall Ball is a function with a Gaussian core and an exponential tail. Here, the central Gaussian is taken to represent the spread of the jet resolution with the exponential tail taking into account events with multiple jets. The mean and width of each peak is determined from the fitting procedure.

The fit means and widths are compared in data and MC in Figure 5.29. The values shown are used to smear and shift the jet p_{T} event by event in the calculation of the

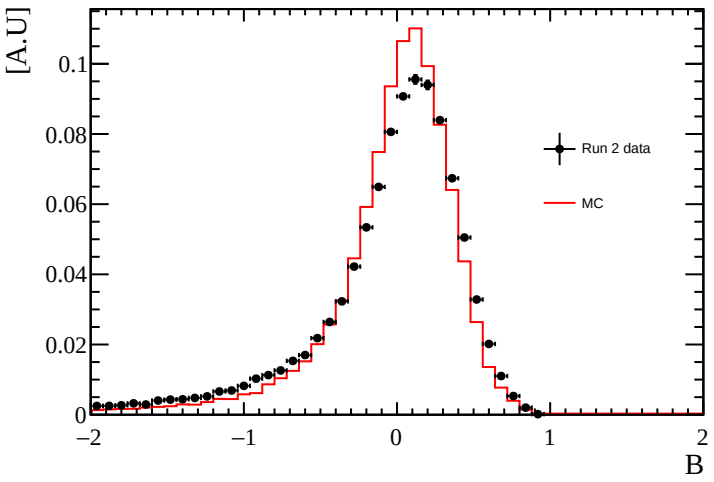

FIGURE 5.28: The balance of the Z boson and jet in $(Z \rightarrow \mu\mu)$ + jet events compared between data and MC.

corrected resolution factor. A plot comparing the balance in data and simulation after the smearing of the MC p_{T} is shown in Figure 5.30.

5.6.2 Electron Momentum Scale

The accuracy with which the resolution factor can be determined also depends on the accuracy with which the electron momentum is reconstructed in data. Specifically, it depends on how well Bremsstrahlung effects are modelled as they smear the p_{T} of the electrons to lower values.

A comparison of the electron momentum scale in data and simulation is investigated. Events are selected in $Z \rightarrow ee$ samples where both electrons must have p_{T} greater than 20 GeV and be inside the LHCb acceptance. The leptons must also pass the identification criteria outlined in Section 5.3.2.3 and be separated by at least 2.7 radians in ϕ. At least one of the electrons in each event must pass all three stages of the trigger. This selection is summarised in Table 5.18.

The agreement between the data and simulation is assessed by defining a global scale factor, k. The p_{T} in simulation is scaled by a range of k factors event by event and the resulting distributions are compared with the data distribution by evaluating the χ^2/nDF agreement for each k factor. The best agreement is seen for a factor of 0.996,

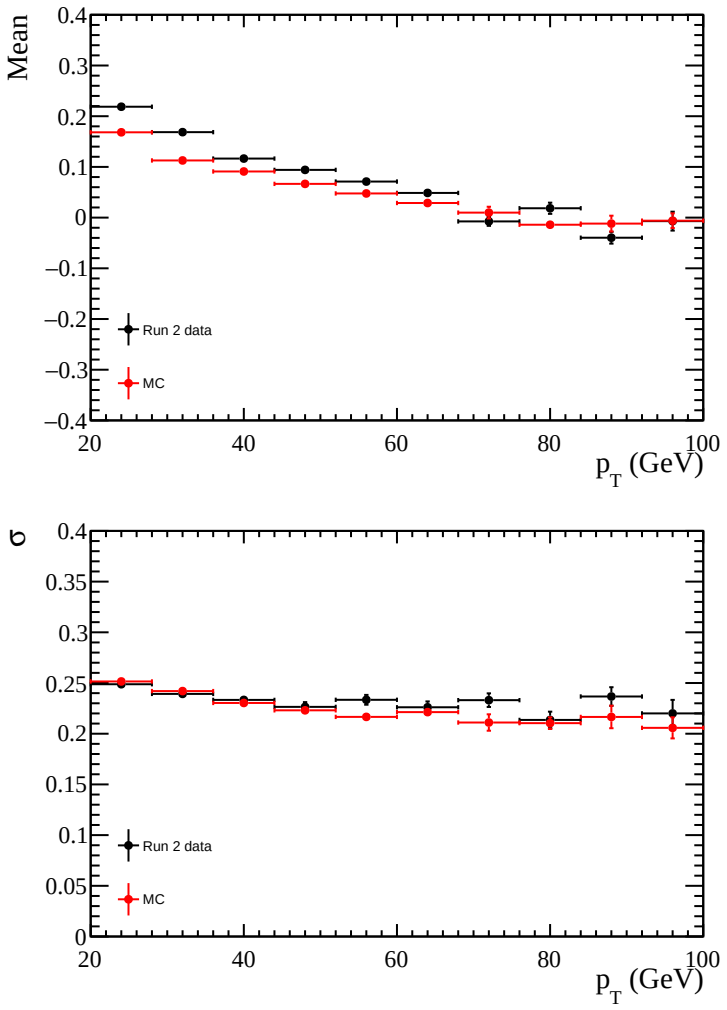

FIGURE 5.29: The mean (top) and resolution (bottom) of the jet balance fits in data and MC as functions of jet p_{T}.

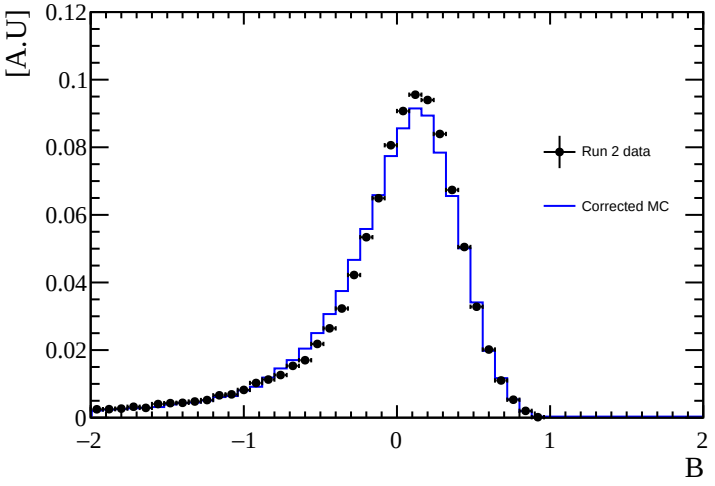

FIGURE 5.30: The balance of the Z boson and jet in $(Z \rightarrow \mu\mu)$ + jet events compared between data and MC after the smearing of the MC p_{T}.

Trigger	`L0Electron`
	`Hlt1SingleElectronNoIP`
	`Hlt2EWSingleElectronVHighPt`
Electrons	`!isMuonLoose`
	$p_{\mathrm{T}} > 20$ GeV, $2.0 < \eta < 4.5$
	$\Delta\phi > 2.7$ rad, $E_{PRS} > 50$ MeV
	$E_{HCAL}/P < 0.05$, $E_{ECAL}/P > 0.1$

TABLE 5.18: A summary of the selection criteria applied to $Z \rightarrow ee$ data and MC to model the electron momentum scale correction to the resolution factor.

as shown in Figure 5.31. This value is used in determining the uncertainty on the resolution factor.

5.6.3 Summary

The central value for the resolution factor is taken from simulation with an NLO correction applied. To assign an uncertainty to the value, the factor is evaluated without the NLO corrections, with the jet energy scale correction applied and with the electron momentum scale correction applied in turn. The difference between these values and the central value are added in quadrature to give the systematic uncertainty. Using this method, the resolution factor is found to be 1.20 ± 0.02.

5.7 Cross-section Calculation

The top pair production cross-section is determined using Equation 5.7,

$$\sigma_{t\bar{t}} = \frac{(N - N_{bkg}) \cdot \mathcal{F}_{res}}{\mathcal{L} \cdot \varepsilon_{sel} \cdot \varepsilon_{tag} \cdot \varepsilon_{rec}}, \tag{5.7}$$

where N is the total number of candidates selected and N_{bkg} is the expected number of background events as described in Section 5.4. The resolution factor is given by $\mathcal{F}_{res}$ and the luminosity by $\mathcal{L}$. The selection, tagging and reconstruction efficiencies are given by ε_{sel}, ε_{tag} and ε_{rec} respectively and are described in the previous sections. Values for each variable are summarised in Table 5.19.

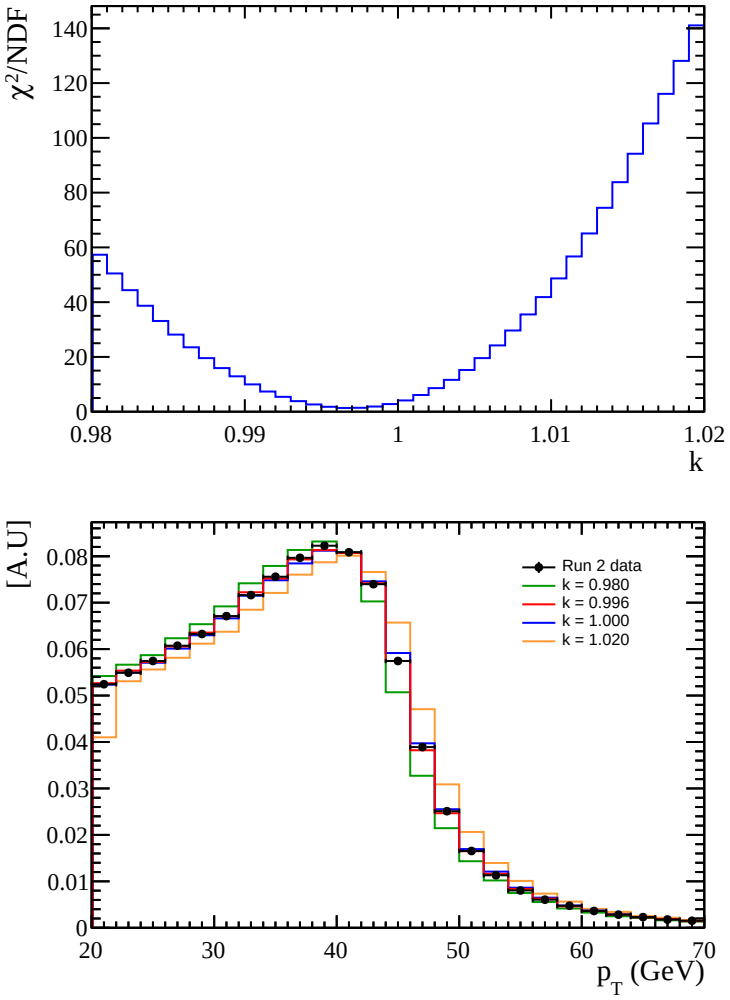

FIGURE 5.31: The reduced χ^2 distribution for a range of scale factors, showing the agreement between $Z \rightarrow ee$ p_{T} distributions in data and simulation (top) and the p_{T} distribution of electrons in $Z \rightarrow ee$ data compared to MC scaled by a sample of k factors (bottom).

N	132±11
N_{bkg}	14.3±2.7
$\mathcal{F}_{res}$	1.20±0.02
$\mathcal{L}$	5.40±0.24
ε_{sel}	0.666±0.020
ε_{tag}	0.556±0.056
ε_{rec}	0.0605±0.039

TABLE 5.19: A summary of the values used to determine the cross-section, as outlined in the previous sections..

5.8 Systematic Uncertainties

A summary of the systematic uncertainties taken into account in this analysis are shown in Table 5.20. The leading systematic on this measurement is seen to be the uncertainty on the jet tagging efficiency. It is expected that this uncertainty will be reduced in future measurements due to the development of a new method for determining the jet tagging efficiency.

Source	Uncertainty (%)
Trigger	1.2
Muon identification	1.2
Muon tracking	2.1
Electron identification	4.8
Electron tracking	2.8
Jet reconstruction	1.3
Jet tagging	10.0
Selection	3.0
Background	2.4
Resolution factor	1.9
Total	12.4

TABLE 5.20: A summary of the systematic uncertainties on the $t\bar{t}$ cross-section measurement.

5.9 Results

The top quark pair production cross-section, $\sigma_{t\bar{t}}$ can be calculated from the inputs described above using Equation 5.7. The cross-section is quoted for the fiducial region ($2.0 < \eta_\ell < 4.5$, $2.2 < \eta_b < 4.2$, $p_{\mathrm{T}}(\ell b) > 20$ GeV), where ℓ refers to the final state leptons and b refers to the b-jet. It is determined to be:

$$\sigma_{t\bar{t}} = 117 \pm 10(\mathrm{stat.}) \pm 15(\mathrm{syst.}) \pm 5(\mathrm{lumi.})\mathrm{pb} \tag{5.8}$$

where the first uncertainty is statistical, the second is systematic and the third is due to the luminosity determination. As this is a fiducial measurement, the result does not rely on the theoretical modelling of the $t\bar{t}$ process and so no signal modelling uncertainties are included. There are two exceptions to this: the resolution factor and the shapes of the

jet and lepton efficiency distributions due to kinematic effects. In both cases, differences between LO and NLO models are considered and found to be negligible.

The measured cross-section can be compared with two theoretical predictions made using POWHEG and aMC@NLO using the NNPDF3.0 PDF set. Three sources of uncertainty are considered for each prediction: the uncertainty due to the description of the PDFs (δ_{PDF}), the uncertainty due to missing higher order corrections (δ_{scale}), and the uncertainty on the strong coupling constant, α_s used in the calculations (δ_{α_s}). All of the uncertainties are evaluated using standard methods. The uncertainty on the PDFs is calculated by taking the variance of the cross-section predictions when each of 100 different replicas of the PDF set is used in turn. The uncertainty due to missing higher order corrections is assessed by varying the renormalisation and factorisation scales by a factor of two around the nominal choice and measuring the resulting difference with respect to the nominal scale. The uncertainty due to the strong coupling constant is included by varying α_s by its uncertainty and recalculating the cross-section.

It is important to note that the scale choices differ slightly between the generators used to calculate the predictions. For POWHEG, the scale choice is taken to be the transverse mass of the final state top quarks. For aMC@NLO, it is taken to be half the sum of the transverse mass of all of the final state particles. These two definitions differ in cases where additional partons are produced in addition to the top quark. No uncertainty is considered from the mass of the top quark.

The uncertainties on the theoretical predictions are combined using Equation 5.9 as recommended by the Higgs cross-section working group [101],

$$\delta_{theory} = \sqrt{\delta^2_{PDF} + \delta^2_{\alpha_s}} + \delta_{scale}, \tag{5.9}$$

to give an overall prediction of $\sigma_{t\bar{t}} = 83^{+18}_{-17}$ pb from POWHEG and $\sigma_{t\bar{t}} = 75^{+13}_{-14}$ pb from aMC@NLO. A comparison of the measured fiducial cross-section and the predictions is shown in Figure 5.32. A good agreement between the measured value and the predictions is observed.

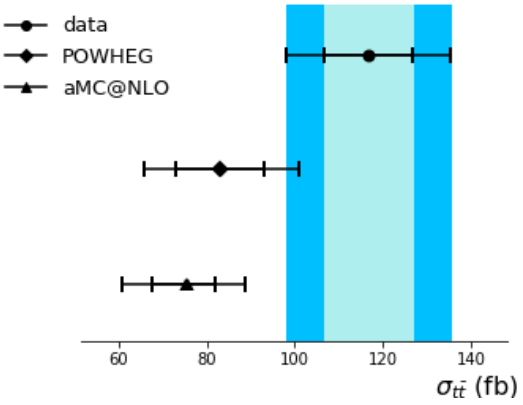

FIGURE 5.32: The measured fiducial cross-section compared with NLO theoretical predications from POWHEG and aMC@NLO. For the data, the inner band represents the statistical uncertainty and the outer band represents the total. For the theoretical predictions, the inner bands represent the scale uncertainty and the outer bands represent the total.

Chapter 6

Feasibility of a WW production cross-section measurement

In this chapter, a feasibility study for the measurement of the WW production cross-section at LHCb with 5.4 fb^{-1} of data is presented. The motivation for the measurement is outlined, followed by a summary of the selection criteria chosen to reduce contamination from various backgrounds and the expected contributions from those backgrounds in the selected sample. A prediction is then made for when the measurement is expected to be feasible.

6.1 Motivation

The measurement of the production rate of W-boson pairs through pp collisions is an important test of the theoretical predictions made by the Standard Model. WW production could also contibute a background to new physics searches.

A LO Feynman diagram of WW production can be seen in Figure 6.1.

The first measurements of WW production were conducted at LEP [102] and then by the CDF [103] and D0 [104] experiments at the Tevatron. At the LHC, the WW production cross-section has been measured by ATLAS and CMS at $\sqrt{s}$ = 7, 8 and 13 TeV [105–111]. The most recent LHC measurements show consistency with NNLO predictions. A summary of the CMS results can be seen in Figure 6.2.

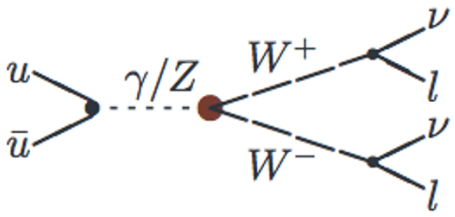

FIGURE 6.1: Feynman diagram of WW production decaying to leptons (l) and neutrinos (ν).

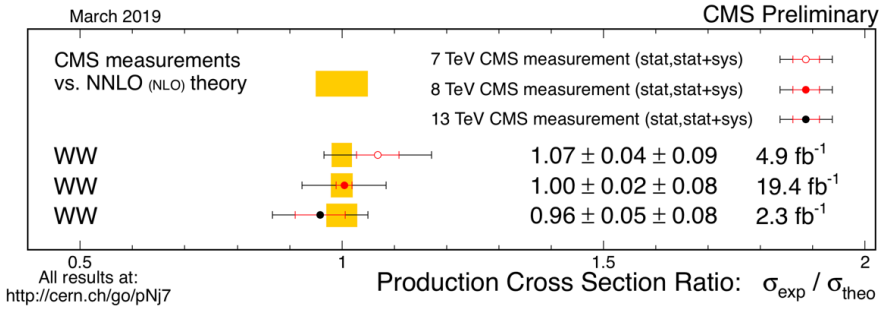

FIGURE 6.2: Summary of WW results measured by CMS during Run-I and Run-II. Adapted from [112].

WW production has not yet been observed by LHCb, due to the low acceptance in Run-I. The principal difficulties in identifying WW events at LHCb are due to the lack of missing momentum information used to identify WW events at the central detectors, and the low jet acceptance for rejecting $t\bar{t}$ background events. However, with the increased cross-section expected from the increase in centre of mass energy for Run-II, and the increased data set expected for Run-II and Run-III, the possibility of making a WW production measurement is now being investigated. The feasibility of measuring the WW production cross-section is assessed below. This analysis focusses on the μe final state, as the choice of different flavour leptons supresses background contamination from $Z \to \mu\mu$ and $Z \to ee$ decays.

6.2 Data and Simulation

As for the top quark pair production measurement, the full Run-II dataset is considered. Data corresponding to a total integrated luminosity of 5.4 fb^{-1} is investigated, as discussed in Section 5.2.

Centrally produced samples of MC simulation are used to model the WW signal and the expected background contributions and their interactions with the detector. These MC samples are produced at LO using Pythia8 and the CT09MCS PDF set. Each sample is scaled by a NLO correction factor determined using aMC@NLO. As for the top quark study, the $t\bar{t}$ MC sample is a weighted combination of events produced through $q\bar{q}$ annihilation and gg fusion.

6.3 Event Selection

As the μe final state is similar to the μeb final state chosen for the $t\bar{t}$ measurement, the selection criteria for isolating the WW signal is similar to that outlined in Section 5.3, with the jet requirements removed. The muon in the event must have passed the `L0EWMuon`, `HLT1SingleMuonHighPT` and `HLT2EWSingleMuonVHighPt` levels of the trigger. Both leptons are again required to have p_{T} greater than 20 GeV and to be within the LHCb acceptance. The muon in the event is required to pass the `isMuon` requirement. The electron is required to pass the high p_{T} electron criteria described in Section 5.3.2.3.

Further requirements are placed on the isolation, impact parameter and vector sum of the lepton p_{T}, as well as a veto on events where high p_{T} jets are present. These are discussed below.

To illustrate the motivation for these requirements, plots comparing distributions of distinguishing variables for signal and background are included. The events in these distributions have been selected with minimal selection criteria: both leptons must have p_{T} greater than 20 GeV and $2.0 < \eta < 4.5$, and the muon must pass all three stages of the trigger.

6.3.1 Isolation

The $p_{\mathrm{T}}^{\mathrm{Cone}}$ variable defined in Section 5.3.2.2 is used to select leptons with isolated tracks to reduce the background contribution from lepton misidentification due to hadronic jets.

Both leptons are required to have a $p_{\mathrm{T}}^{\mathrm{Cone}}$ less than 2 GeV. This value is chosen by investigating the relationship between signal efficiency and purity when combining varying requirements placed on $p_{\mathrm{T}}^{\mathrm{Cone}}$ and ip. The requirements placed on these variables are chosen together to produce the selection with the highest value of efficiency/purity, to ensure the sample is as pure as possible without sacrificing statistical significance.

The $p_{\mathrm{T}}^{\mathrm{Cone}}$ distributions for the leptons can be seen in Figure 6.3, along with the distributions from the same-sign data sample which is used to model the lepton misidentification background as in Section 6.4. Here, the requirement placed at 2 GeV reduces the predicted background contribution from ∼70% to ∼20%.

6.3.2 Impact Parameter

A requirement is placed on the impact parameter to reduce the significant background contribution expected from $Z \rightarrow \tau\tau$ events, similar to that discussed in Section 5.3.2.1. The requirement is slightly tighter than for the $t\bar{t}$ analysis, at 0.035 mm. Impact parameter distributions of WW and $Z \rightarrow \tau\tau$ simulation are compared in Figure 6.4, where 65% of the $Z \rightarrow \tau\tau$ events are removed by the ip requirement and 90% of the WW events are retained.

6.3.3 Jet Multiplicity

A requirement on jet multiplicity is used to remove background contribution from $t\bar{t}$ events. Top pair production events contain b-jets and WW events are less likely to include jets, so the jet multiplicity can be used as a distinguishing variable. Events are required to have no jets with p_{T} greater than 20 GeV. Plots showing the number of jets present in simulated WW and $t\bar{t}$ events can be compared in Figure 6.5. The jet veto removes ∼60% of the $t\bar{t}$ events and retains ∼90% of the WW events.

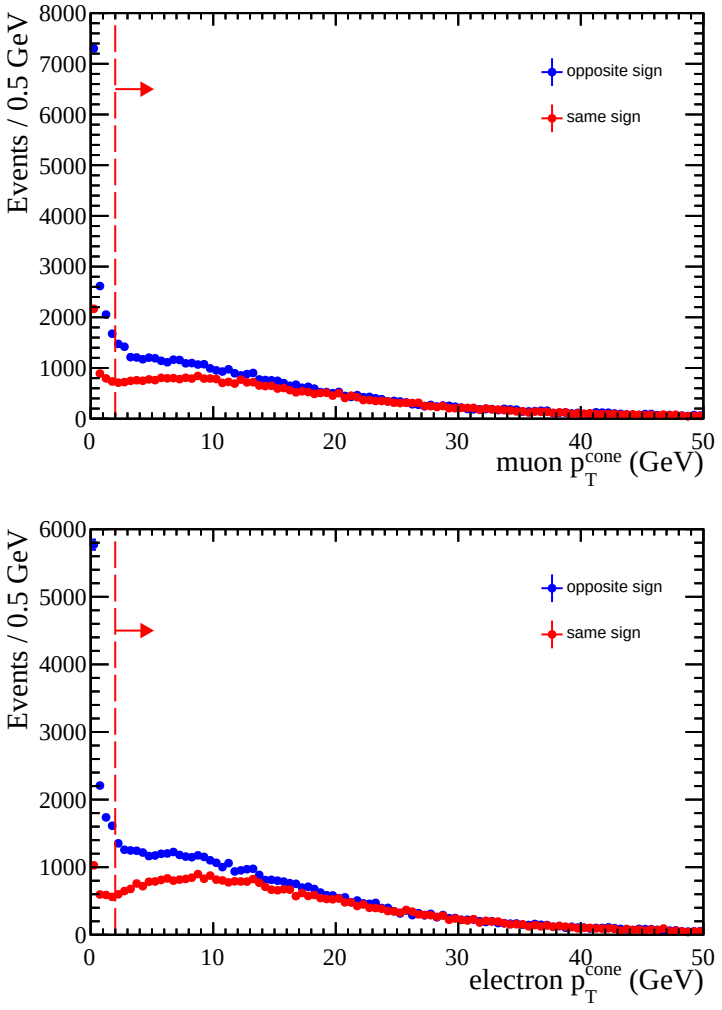

FIGURE 6.3: Comparisons of $p_{\mathrm{T}}^{\mathrm{Cone}}$ distributions in opposite- and same-sign data in blue and red respectively, for muons (top) and electrons (bottom). The red dashed line shows the position of the $p_{\mathrm{T}}^{\mathrm{Cone}}$ requirement with the arrow indicating which events are removed.

6.3.4 p_{T} Vector Sum

At ATLAS and CMS, the presence of neutrinos can be inferred from missing momentum and momentum in the reconstructed events. This is not possible at LHCb due to its limited forward acceptance. Instead, the vector sum of the lepton transverse momentum, Σp_{T}, can be used to distinguish between WW and $Z \to \tau\tau$ events. As there are more neutrinos in the $Z \to \tau\tau$ final state than the WW, more p_{T} will be carried away by undetected particles and the Σp_{T} of the remaining leptons will be lower. Distributions of the p_{T} vector sum are compared for simulated WW and $Z \to \tau\tau$ events in Figure 6.6.

The Σp_{T} must be greater than 60 GeV. The requirement placed on this variable is chosen

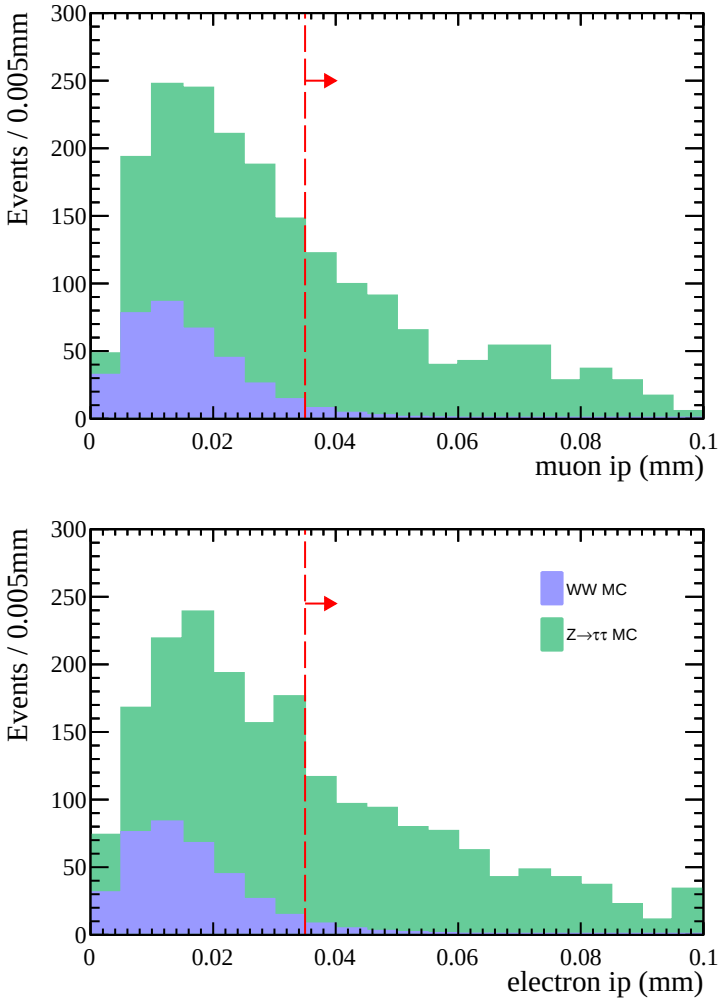

FIGURE 6.4: Overlaid comparisons of WW and $Z \to \tau\tau$ MC ip distributions, in blue and green respectively, for the muon (top) and electron (bottom). The distributions are normalised to 5.4 fb^{-1}, with the red line showing the placement of the ip requirement with the arrow indicating which events are rejected.

to remove ∼95% of the $Z \to \tau\tau$ background contamination and retains ∼40% of the WW signal. In future studies this requirement will be optimised to remove as many $Z \to \tau\tau$ events as possible while retaining a significant signal contribution by comparing the signal efficiency and purity for a range of Σp_{T} thresholds.

6.3.5 Summary

The selection criteria used to isolate the WW signal is summarised in Table 6.1. A total of 223 events are selected.

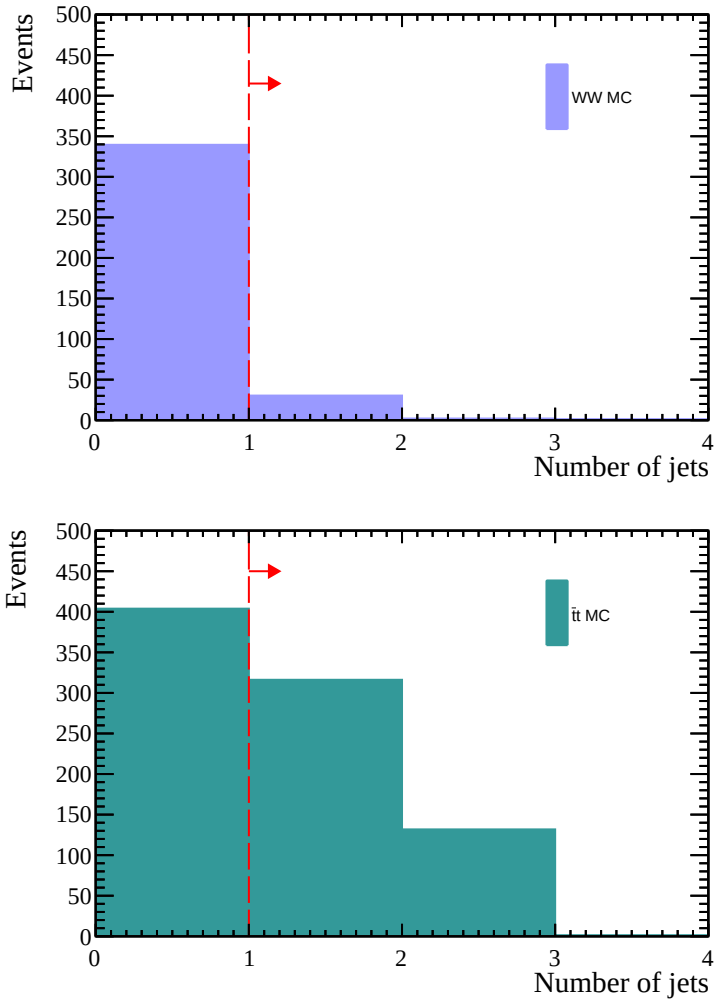

FIGURE 6.5: The jet multiplicity in WW (top) and $t\bar{t}$ (bottom) MC distributions, normalised to 5.4 fb^{-1}. The red lines show the requirement used in the event selection with the arrow indicating which events are rejected.

Trigger	`L0EWMuon`
	`HLT1SingleMuonHighPT`
	`HLT2EWSingleMuonVHighPt`

Muon	`isMuon` $p_{\mathrm{T}} > 20$ GeV, $2.0 < \eta < 4.5$ $ip < 0.035$ mm, $p_{\mathrm{T}}^{\mathrm{Cone}} < 2$ GeV
Electron	$E_{PRS} > 50$ MeV, $E_{HCAL}/P < 0.05$ $E_{ECAL}/P > 0.1$, `!isMuonLoose` $p_{\mathrm{T}} > 20$ GeV, $2.0 < \eta < 4.5$ $ip < 0.035$ mm, $p_{\mathrm{T}}^{\mathrm{Cone}} < 2$ GeV
Leptons	$\Sigma p_{\mathrm{T}} > 60$ GeV

TABLE 6.1: A summary of the selection criteria applied to isolate the WW signal.

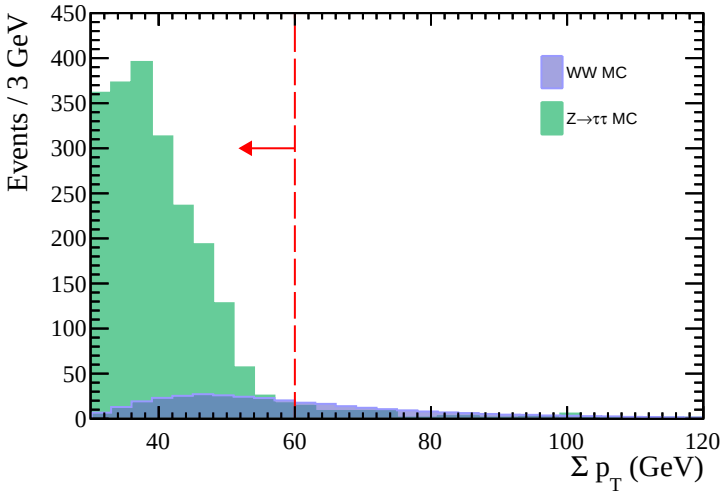

FIGURE 6.6: Overlaid comparison of WW and $Z \to \tau\tau$ MC Σp_{T} distributions, in blue and green respectively. Both distributions are normalised to 5.4fb^{-1}. The red line shows the position of the requirement with the arrow showing which reqion is rejected.

6.4 Signal Purity

Using the selection criteria outlined in Section 6.3, 223 events are selected. To determine the signal purity of the sample, contributions from a number of backgrounds are considered using models taken from both MC and data.

The lepton misidentification background is modelled with same-sign data as for the $t\bar{t}$ measurement. The technique used to determine the expected number of lepton misidentification events in the sample is the same as that described in Section 5.4.1. Here, the number of events expected in signal Region A is determined by applying a scaling factor to the events selected in control Region B, where the full WW selection criteria is applied to events with same-sign leptons. The scaling factor is determined as the ratio of events in Regions C and D, where the same selection but with reversed electron identification requirements is applied to opposite- and same-sign events respectively. For this analysis, the scale factor is determined to be 2.93 and the expected contribution from lepton misidentification is 58.2±12.7 events, where the uncertainty is taken from the statistical uncertainty in each control region added in quadrature.

The other backgrounds being considered are all modelled using simulation. Their contibutions are summarised in Table 6.2. The backgrounds are:

- $Z \to \tau\tau$ events where the taus decay to a muon and electron final state,

- $Z \rightarrow \mu\mu$ events where one of the muons is misidentified as an electron due to deposits in the calorimeter system,
- $Z \rightarrow ee$ events where one of the electrons is misidentified as a muon, having punched through to the muon stations,
- $t\bar{t}$ decays to the μe channel,
- $(W \rightarrow \mu\nu)$ + jet production where the jet is misidentified as an electron.

The uncertainty assigned to each of these backgrounds is the expected statistical uncertainty on the number of predicted events and uncertainty due to the limited size of the MC samples added in quadrature. For the latter, the largest uncertainties come from the backgrounds with the smallest number of MC events available.

The potential contributions from other diboson backgrounds are not included as no full simulation of such events is available. The contributions are expected to be significantly smaller than the expected WW signal, but may be considered further when more statistics are available.

The expected contribution from each background is given in Table 6.2. From simulation, the expected number of signal events is 61.3±0.6.

Source	Method	Events
Lepton mis-identification	Same-sign data	58.2±12.7
$Z \rightarrow \tau\tau$	Simulation	11.4±6.9
$t\bar{t}$	Simulation	37.4±6.2
$Z \rightarrow ee$	Simulation	3.9±3.5
$(W \rightarrow \mu\nu)$ + jet	Simulation	6.5±3.6
$Z \rightarrow \mu\mu$	Simulation	6.3±3.2
Total		123.8±27.4

TABLE 6.2: A summary of the expected contributions from each background and the methods used to evaluate them.

6.5 Results

Comparisons of the data and expected backgrounds can be seen in Figures 6.7, 6.8 and 6.9 for distributions of the μe invariant mass, lepton transverse momenta and lepton pseudorapidities.

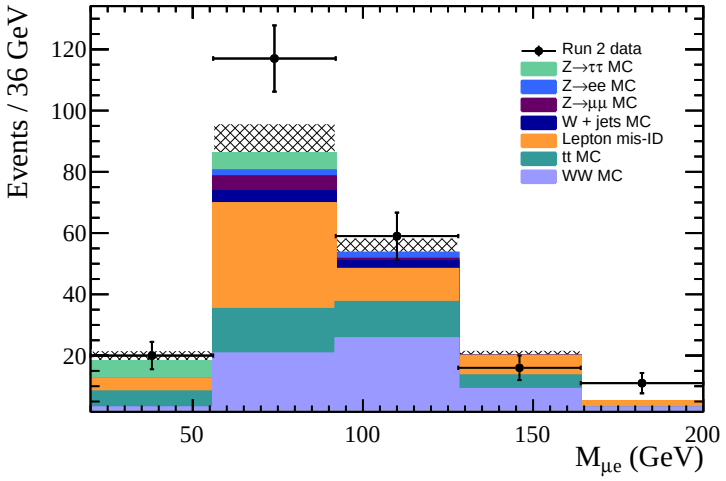

FIGURE 6.7: Invariant mass of the combined muon and electron, comparing data and expected model expected contributions. The grey hashed areas represent the total uncertainty from all of the contributions in the model. The final bin is dominated by signal and so the background uncertainty is too small to display

6.6 Outlook

A typical method for assessing the feasibility of a measurement is to evaluate the significance, S. This is an estimate of the precision with which a measurement can be verified when compared to its associated uncertainty. Typically, in a counting measurement such as this, the significance can be determined from the number of expected signal events, N_{signal}, and the number of expected background events, N_{bkg}, using:

$$S = \frac{N_{signal}}{\sqrt{N_{signal} + N_{bkg}}} \tag{6.1}$$

A standard evidence result is usually quoted with a significance of 3σ. For an observation or measurement the significance is usually required to be 5σ. For this analysis, the value of S is 4σ. However this does not take into account the precision of the model used to estimate the signal and background contributions and is therefore an over-estimate of the significance of this result. Instead, an expanded definition of significance, S' is used:

$$S' = \frac{N_{signal}}{\sqrt{N_{signal} + N_{bkg} + \sigma^2_{MC} + \sigma^2_{DD}}} \tag{6.2}$$

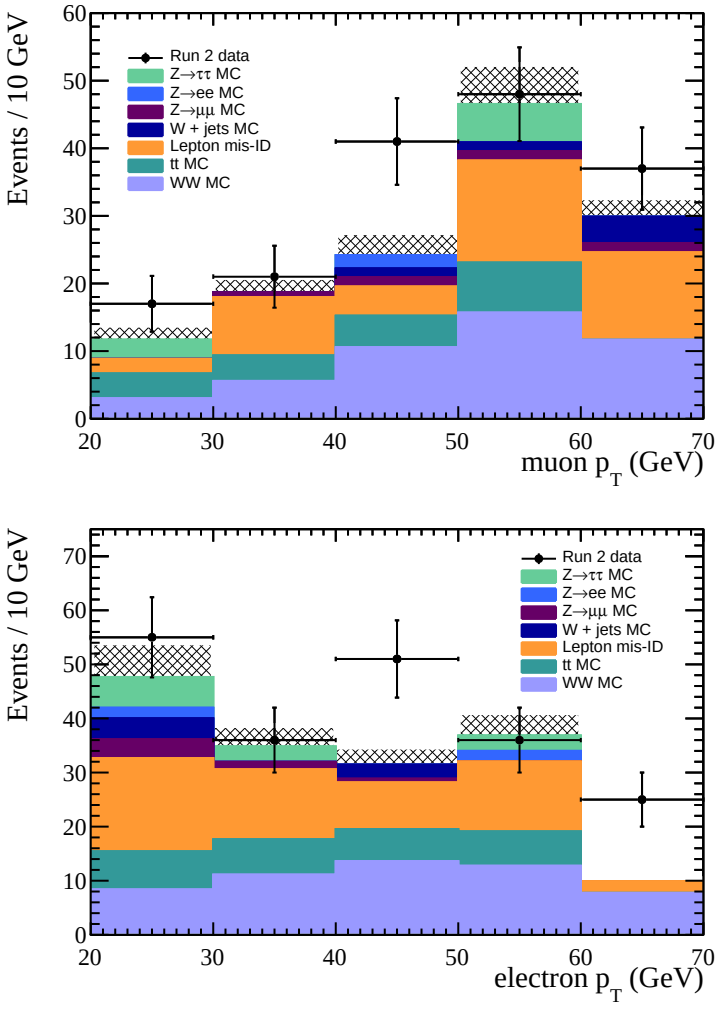

FIGURE 6.8: The transverse momentum distribution for muons (top) and electrons (bottom) in data compared with the expected contributions. The grey hashed areas represent the total uncertainty from all of the contributions in the model. The final bin in the electron plot is dominated by signal and so the background uncertainty is too small to display.

where σ_{MC} and σ_{DD} are the statistical uncertanties on the simulated and data-driven portions of the background model respectively. The σ_{MC} values for the two leading MC modelled backgrounds ($t\bar{t}$ and $Z \rightarrow \tau\tau$) also include an additional uncertainty to account for the precision with which their contributions are calculated. For the $t\bar{t}$, this is the uncertainty on the aMC@NLO prediction discussed in Section 5.9 and for $Z \rightarrow \tau\tau$ it is taken from the total uncertainty on the final di-lepton measurement in Reference [113]. For this analysis, the value of S' is 2.8σ.

Using S', it is possible to estimate how much data would be necessary to make a measurement of the WW cross-section by introducing a data scale factor and projecting the increase in significance as the available data set grows.

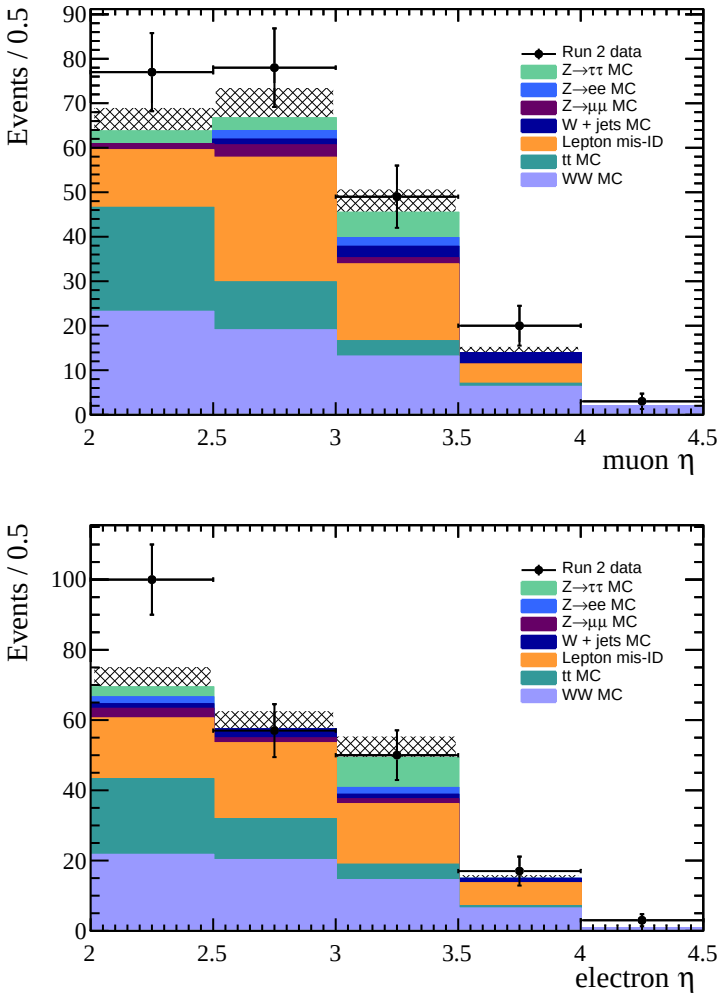

FIGURE 6.9: The pseudorapidity distribution for muons (top) and electrons (bottom) in data compared with the expected contributions. The grey hashed areas represent the total uncertainty from all of the contributions in the model. The final bin in each plot is dominated by signal and so the background uncertainties are too small to display.

As the data available for measuring the cross-section increases, so will the number of events in the control regions used to calculate the data-driven background contribution and its uncertainty. It is also reasonable to assume that the available centrally produced MC will scale with the data. With all of this included, the estimated scale factor needed to reach 5σ is 3.5. This means that an integrated luminosity of approximately 19 fb^{-1} would be needed to measure the WW production cross-section, which is expected to be recorded during Run-III as can be seen in Figure 6.10.

It is possible to make an additional assumption regarding the modelling of the MC background. As centrally produced MC samples must be suitable for use in a wide range of physics analyses, the samples are not optimised for this particular measurement

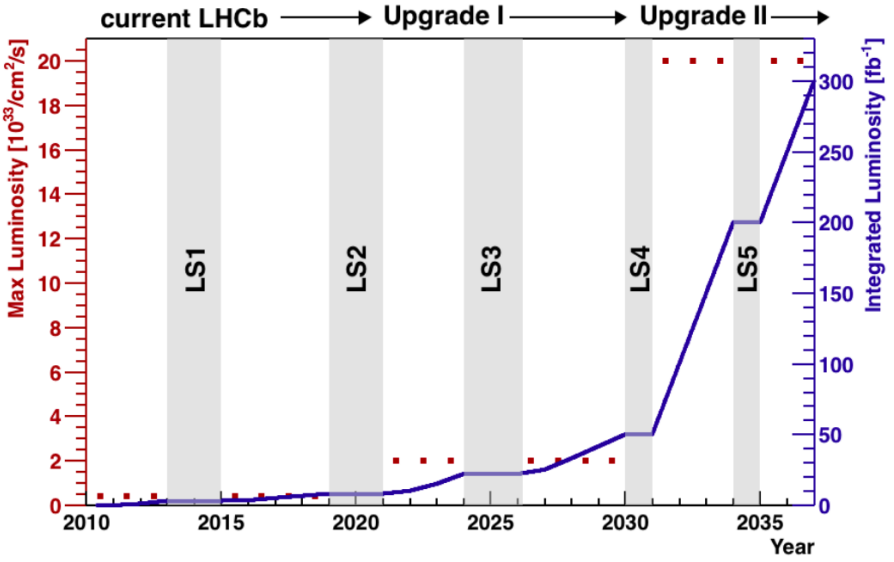

FIGURE 6.10: Luminosity projections for LHCb upgrades as a funtion of time. The red points and left scale indicate the anticipated instantaneous luminosity during each period while the blue line and right scale indicate the total integrated luminosity [86].

and so some have very low statistics which contribute to a large uncertainty on the MC model. It would be possible to improve this uncertainty by producing dedicated samples of $Z \to \tau\tau$, $Z \to \mu\mu$, $Z \to ee$ and $(W \to \mu\nu)$ + jet MC, where requirements are introduced at the generator level. Using this method, it is generally possible to increase the statistics in a simulated sample by around a factor of 10. If it is assumed that the available simulation of all four of these samples can be increased by this factor, the data scale factor needed to reach an S' of 5σ is reduced slightly to 3. This is the equivalent of a total integrated luminosity of $\sim$16 fb^{-1}.

With an increased data set, it would also be possible to make further improvements to the method used to isolate the WW signal. It is possible that certain requirements in the selection criteria could be tightened to improve purity as the signal efficiency would not be as much of a priority with more events available. It is also possible that the measurement could benefit from the application of multi-variate analysis with more available data and improved MC statistics. This would help to optimise the selection criteria and result in a sample with a higher purity, again improving the significance of the measurement.

Chapter 7

Conclusion

This thesis presents methods for testing the theoretical predictions made by the Standard Model by measuring the production cross-sections of particles produced at the Large Hadron Collider.

A measurement of the top quark pair production cross-section in the μeb final state in the forward region is evaluated using 5.4 fb^{-1} of data collected by the LHCb experiment during LHC Run-II at $\sqrt{s}$=13 TeV. The muon, electron and b-jet in each event are subject to selection criteria chosen to maximise the purity of the sample. Contributions from a number of backgrounds are considered using Monte Carlo simulation and data-driven techniques. The efficiency with which events are reconstructed and selected is evaluated and used to calculate a fiducial cross-section of:

$$\sigma_{t\bar{t}} = 117 \pm 10(\text{stat.}) \pm 15(\text{syst.}) \pm 5(\text{lumi.})\text{pb} \tag{7.1}$$

where the first uncertainty is statistical, the second is systematic and the third is due to the luminosity determination. Comparisons are made with theoretical predictions calculated from POWHEG and aMC@NLO and the measurement is seen to be in good agreement.

The feasibility of a WW production cross-section measurement with the same data set is also investigated. It is concluded that such a measurement is possible at LHCb, but not with the currently available statistics. Using an estimate of the expected sensitivity, a prediction of when the measurement will be possible can be made. It is estimated that

it will be possible to make the necessary improvements to MC and collect sufficient data to measure the WW production cross-section in the μe channel with a significance of 5σ within the next 3 years of data taking.

List of Figures

2.1 The shape of the Higgs potential for $\mu^2 < 0$. The indicated motion of the blue ball illustrates Spontaneous Symmetry Breaking which results in a non-zero Vacuum Expectation Value for the Higgs field [9]. 10
2.2 A CKMfitter Collaboration fit to the Unitarity Triangle with individual constraints from various measurements superimposed [11]. 15
2.3 A summary of $t\bar{t}$ production cross-section measurements at the ATLAS and CMS experiments [25]. 19
2.4 The production ratio of top quarks as a function of top pseudorapidity at $\sqrt{s}$ 14 TeV [27]. 20
2.5 Feynman diagrams for the production of top quark pairs in leading order processes at the LHC: (a) gluon fusion, (b) pair creation from $q\bar{q}$ annihilation, (c) t-channel gluon fusion. 21
2.6 Examples of next-to-leading order production meachanisms for top quark pair production. 21
2.7 The different branching ratios of a top quark pair, referring to the decay modes of the two W bosons arising from the top quark decays [28]. 22

3.1 The LHC ring with the CERN site and four experiments labelled [29]. . . 24
3.2 The LHC accelerator complex [30]. 25
3.3 The instantaneous luminosity over time at ATLAS, CMS and LHCb during fill 2651 of the LHC in May 2012 [35]. 27
3.4 The integrated luminosity recorded by LHCb per year during Run-I and Run-II [36]. 27
3.5 The angular direction of $b\bar{b}$ production in simulated collisions at 14 TeV, with the LHCb acceptance indicated in red [40]. 28
3.6 The LHCb experiment with each detector sub-system labelled. Particles travel from left to right through the detector [41]. 29
3.7 Layout of LHCb tracking stations. The TT is on the left of the image, T1-T3 are on the right. Silicon trackers are displayed in purple and drift-time trackers in blue. The VELO is situated downstream of the TT station at a lower position on the z axis [41]. 31
3.8 The layout of the LHCb VELO showing the position of the 42 modules around the interaction region. Fully closed and fully open module positions are also shown [41]. 32
3.9 Detailed view of a pair of closed VELO modules with R and ϕ geometry [41]. 32
3.10 Schematic of LHCb's dipole magnet [41]. 33
3.11 The (x, u), (v, x) configuration of silicon sensor layers in the TT [46]. . . . 34

3.12 Layout of an x (top) and u (bottom) layer in the IT with each box representing a sensor [47]. . . . 35
3.13 Schematic of the OT in T1-T3. Cross-section of a detector module with measurements in mm (a) and configuration of the modules (b) [48]. . . . 36
3.14 The reconstructed Cherenkov angle in RICH-1 as a function of track momentum. Distinct curves have been labelled for a variety of particles [49]. 37
3.15 The calorimeter system [51]. . . . 38
3.16 Cell structure in the ECAL [41]. . . . 39
3.17 Cell structure in the HCAL [41]. . . . 39
3.18 The layout of the muon stations with the calorimeter system and iron filters in between [41]. . . . 41
3.19 The passage of muons through matter, showing the mean energy loss per distance travelled [1]. . . . 42

4.1 Different type of tracks [55]. . . . 45
4.2 A graphical representation of the different interactions of particles throughout the layers of a standard particle physics experiment [62]. . . . 48
4.3 A summary of the Run-II trigger scheme with data outputs at each level displayed [68]. . . . 51
4.4 A summary of LHCb software chain. [83]. . . . 56

5.1 The kinematics of LHC experiments at 13 TeV [84]. . . . 58
5.2 Graphical representation of a pp collision at LHCb. For the particles produced in the interaction to enter LHCb acceptance, parton 1 must have a high Bjorken-x while the Bjorken-x of parton 2 must be low for the momentum transfer to boost the final state particles into the forward region [84]. . . . 58
5.3 Feynman diagram of top quark pair production through gluon fusion decaying to leptons (ℓ) and neutrinos (ν). . . . 59
5.4 Predicted reduction in gluon PDF uncertainty with the inclusion of an LHCb $t\bar{t}$ cross-section measurement with uncertainty of 4-8% [27]. . . . 60
5.5 An overlay of impact parameter distributions to compare $t\bar{t}$ and $Z \to \tau\tau \to \mu e$ MC for the muon (top) and electron (bottom). The distributions are normalised to 5.4 fb^{-1}. The dashed red lines shows the position of the ip requirement with the arrow indicating which events are rejected. . . . 63
5.6 Comparisons of $p_{\mathrm{T}}^{\mathrm{Cone}}$ distributions in opposite- and same-sign data in blue and red respectively, for muons (top) and electrons (bottom). The red dashed line shows the position of the $p_{\mathrm{T}}^{\mathrm{Cone}}$ requirement with the arrow showing the region being rejected. . . . 65
5.7 A graphical representation of the signal region, A, and three control regions, B, C and D, used to evaluate the contribution from the lepton misidentification background. . . . 68
5.8 Feynman diagram of a Z + jet event. . . . 69
5.9 Invariant mass distribution of the $(Z \to \mu\mu)$ + jet events selected from the full Run-II dataset. These events are used to predict the Z + jet background contribution. . . . 70
5.10 Feynman diagram of a single top event. . . . 70
5.11 The muon p_{T} (top) and pseudorapidity (bottom) distributions in data compared to the expected background contributions. . . . 72

5.12 The electron p_{T} (top) and pseudorapidity (bottom) distributions in data compared to the expected background contributions. . . . 73
5.13 The jet p_{T} (top) and pseudorapidity (bottom) distributions in data compared to the expected background contributions. . . . 74
5.14 The invariant mass of the muon, electron and b-jet in data compared to the expected background contributions. . . . 74
5.15 The muon trigger efficiency as a function of p_{T} (top) and η (bottom) for data and simulation. . . . 76
5.16 The muon identification efficiency as a function of p_{T} (top) and η (bottom) for data and simulation. . . . 78
5.17 The muon tracking efficiency as a function of p_{T} (top) and η (bottom) for data and simulation. . . . 80
5.18 Plots to represent the same-sign data subtraction from opposite-sign events while evaluating the electron identification efficiency in electron p_{T} (top) and pseudorapidity (bottom) distributions. . . . 82
5.19 The electron identification efficiency as a function of p_{T} (top) and η (bottom) for data and simulation. . . . 83
5.20 Template fits to $Z \to e\gamma$ (top) and $Z \to ee$ (bottom) data used in evaluating the electron tracking efficiency. . . . 85
5.21 Template fits $Z \to e\gamma$ (top) and $Z \to ee$ (bottom) MC used in evaluating the electron tracking efficiency. . . . 86
5.22 Distributions of jet quality variables used in assessing the uncertainty on the jet reconstruction efficiency: the charged particle fraction of the jet (top) and p_{T} of the main particle in the jet (bottom). . . . 88
5.23 Distributions of jet quality variables used in assessing the uncertainty on the jet reconstruction efficiency: the fraction of the jet energy carried by the main particle in the jet (top) and the number of tracks with information pointing to the primary vertex (bottom). . . . 89
5.24 Tagging variables investigated for jet tagging efficiency: the minimum flight distance (top) and flight distance χ^2 (bottom). . . . 91
5.25 Tagging variables investigated for jet tagging efficiency: SV lifetime (top) and z position (bottom). . . . 92
5.26 The impact parameter distribution of the muons in $Z \to \mu\mu$ events compared between data, MC and tuned MC. . . . 93
5.27 The impact parameter distributions before and after tuning for electrons (top) and muons (bottom) in $t\bar{t}$ MC. . . . 94
5.28 The balance of the Z boson and jet in $(Z \to \mu\mu)$ + jet events compared between data and MC. . . . 97
5.29 The mean (top) and resolution (bottom) of the jet balance fits in data and MC as functions of jet p_{T}. . . . 98
5.30 The balance of the Z boson and jet in $(Z \to \mu\mu)$ + jet events compared between data and MC after the smearing of the MC p_{T}. . . . 98
5.31 The reduced χ^2 distribution for a range of scale factors, showing the agreement between $Z \to ee$ p_{T} distributions in data and simulation (top) and the p_{T} distribution of electrons in $Z \to ee$ data compared to MC scaled by a sample of k factors (bottom). . . . 100

5.32 The measured fiducial cross-section compared with NLO theoretical predications from POWHEG and aMC@NLO. For the data, the inner band represents the statistical uncertainty and the outer band represents the total. For the theoretical predictions, the inner bands represent the scale uncertainty and the outer bands represent the total. 103

6.1 Feynman diagram of WW production decaying to leptons (l) and neutrinos (ν). 106
6.2 Summary of WW results measured by CMS during Run-I and Run-II. Adapted from [112]. 106
6.3 Comparisons of $p_{\mathrm{T}}^{\mathrm{Cone}}$ distributions in opposite- and same-sign data in blue and red respectively, for muons (top) and electrons (bottom). The red dashed line shows the position of the $p_{\mathrm{T}}^{\mathrm{Cone}}$ requirement with the arrow indicating which events are removed. 109
6.4 Overlaid comparisons of WW and $Z \to \tau\tau$ MC ip distributions, in blue and green respectively, for the muon (top) and electron (bottom). The distributions are normalised to 5.4 fb^{-1}, with the red line showing the placement of the ip requirement with the arrow indicating which events are rejected. 110
6.5 The jet multiplicity in WW (top) and $t\bar{t}$ (bottom) MC distributions, normalised to 5.4 fb^{-1}. The red lines show the requirement used in the event selection with the arrow indicating which events are rejected. 111
6.6 Overlaid comparison of WW and $Z \to \tau\tau$ MC Σp_{T} distributions, in blue and green respectively. Both distributions are normalised to 5.4fb^{-1}. The red line shows the position of the requirement with the arrow showing which reqion is rejected. 112
6.7 Invariant mass of the combined muon and electron, comparing data and expected model expected contributions. The grey hashed areas represent the total uncertainty from all of the contributions in the model. The final bin is dominated by signal and so the background uncertainty is too small to display 114
6.8 The transverse momentum distribution for muons (top) and electrons (bottom) in data compared with the expected contributions. The grey hashed areas represent the total uncertainty from all of the contributions in the model. The final bin in the electron plot is dominated by signal and so the background uncertainty is too small to display. 115
6.9 The pseudorapidity distribution for muons (top) and electrons (bottom) in data compared with the expected contributions. The grey hashed areas represent the total uncertainty from all of the contributions in the model. The final bin in each plot is dominated by signal and so the background uncertainties are too small to display. 116
6.10 Luminosity projections for LHCb upgrades as a funtion of time. The red points and left scale indicate the anticipated instantaneous luminosity during each period while the blue line and right scale indicate the total integrated luminosity [86]. 117

List of Tables

2.1 The properties of the quarks in the Standard Model [1]. 4
2.2 The properties of the leptons in the Standard Model [1]. 4
2.3 The properties of the bosons in the Standard Model [1]. 5

4.1 Required hits in muon stations M2-M5 for a particle to be fully identified as a muon, satisfying the binary `isMuon` criterion [63]. 49
4.2 Required hits in muon stations M2-M5 for a particle to be loosely identified as a muon, satisfying the binary `isMuonLoose` criterion [63]. 49
4.3 Examples of Run-II L0 trigger thesholds with n_{SPD} referring to the multiplicity at the SPD sub-detector. 53
4.4 Examples of Run-II HLT1 trigger lines. The `L0Req` requirement is used to select events passing the given L0 line. The `TrChi2` variable allows for quality cuts on track fits. The `MaxOTHits`, `MaxITHits` and `MaxVeloHits` requirements are Global Event Cuts limiting the maximum multiplicity of the each event in the OT, IT and VELO respectively. 54
4.5 Examples of Run-II HLT2 trigger lines. The `Input` is the location from which the trigger line should select the required particle candidates. The `L0Req` and `Hlt1Req` are used to only select events that have passed the given L0 and HLT1 lines respectively. The `PrsMin`, `EcalMin` and `HCALMax` variables define the minimum or maximum deposits required in the PRS, ECAL and HCAL respectively. `TkChi2` defines the cut on the minimum quality of the track fit. 55

5.1 The predicted production cross-sections of $t\bar{t}$ final states at 7, 8 and 14 TeV within LHCb acceptance using POWHEG predictions matched to Pythia8. The uncertainties include contributions from scale, PDF, showering and jet tagging [85]. 59
5.2 The mean Bjorken-x of the most energetic parton probed by different partially reconstructed $t\bar{t}$ final states [86]. 60
5.3 The integrated luminosity recorded by LHCb for each year in Run-II as used in this analysis. 61
5.4 A summary of the selection criteria applied to isolate $t\bar{t}$ signal. 66
5.5 A summary of the selection criteria applied to isolate $(Z \to \mu\mu)$ + jet candidates. 69
5.6 A summary of the expected contributions from each background and the methods used to evaluate them. 72
5.7 A summary of the selection criteria applied to $Z \to \mu\mu$ data to calculate the trigger efficiency. 76
5.8 A summary of the selection criteria applied to $Z \to \mu\mu$ data to calculate the muon identification efficiency. 77

5.9 A summary of the selection criteria applied to $Z \rightarrow \mu\mu$ data to calculate the muon tracking efficiency. 79
5.10 A summary of the selection criteria applied to $Z \rightarrow ee$ data to calculate the electron identification efficiency. 81
5.11 A summary of the selection criteria applied to $Z \rightarrow e\gamma$ events to calculate the electron tracking efficiency. 83
5.12 A summary of the selection criteria applied to $Z \rightarrow ee$ events to calculate the electron tracking efficiency. 84
5.13 A summary of the selection criteria applied to isolate $(Z \rightarrow \mu\mu)$ + jet candidates to evaluate the uncertainty on the jet reconstruction efficiency. 86
5.14 A summary of the selection criteria applied to $(W \rightarrow \mu\nu)$ + jet data to investigate the jet tagging efficiency. 90
5.15 A summary of the selection criteria applied to isolate $Z \rightarrow \mu\mu$ data and simulation to validate the impact parameter tuning required to evaluate the selection efficiency. 93
5.16 A summary of the efficiency to reconstruct, select and tag $t\bar{t} \rightarrow \mu eb$ events. 95
5.17 A summary of the selection criteria applied to $(Z \rightarrow \mu\mu)$ + jet data to calculate the jet balance. 96
5.18 A summary of the selection criteria applied to $Z \rightarrow ee$ data and MC to model the electron momentum scale correction to the resolution factor. . . 99
5.19 A summary of the values used to determine the cross-section, as outlined in the previous sections.. 100
5.20 A summary of the systematic uncertainties on the $t\bar{t}$ cross-section measurement. 101

6.1 A summary of the selection criteria applied to isolate the WW signal. . . 111
6.2 A summary of the expected contributions from each background and the methods used to evaluate them. 113

Bibliography

[1] M. Tanabashi et al. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018. doi: 10.1103/PhysRevD.98.030001. URL https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[2] P.A.M Dirac. "the quantum theory of the electron". *Proceedings of the Royal Society of London a: Mathematical, Physical and Engineering Sciences*, 117(778): 610–624, 1928.

[3] O. Johns and O.D. Johns. *Analytical Mechanics for Relativity and Quantum Mechanics*. Oxford Graduate Texts. Oxford University Press, 2016. ISBN 9780198766803. URL https://books.google.co.uk/books?id=wQIHswEACAAJ.

[4] David J. Gross and Frank Wilczek. Ultraviolet Behavior of Nonabelian Gauge Theories. *Phys. Rev. Lett.*, 30:1343–1346, 1973. doi: 10.1103/PhysRevLett.30.1343. [,271(1973)].

[5] M. Gell-Mann. The interpretation of the new particles as displaced charge multiplets. *Nuovo Cim.*, 4(S2):848–866, 1956. doi: 10.1007/BF02748000.

[6] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964. doi: 10.1103/PhysRevLett.13.508. [,160(1964)].

[7] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.

[8] Serguei Chatrchyan et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.

[9] John Ellis. Higgs Physics. (KCL-PH-TH-2013-49. KCL-PH-TH-2013-49. LCTS-2013-36. CERN-PH-TH-2013-315):117–168. 52 p, Dec 2013. doi: 10.5170/CERN-2015-004.117. URL https://cds.cern.ch/record/1638469. 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report.

[10] Makoto Kobayashi and Toshihide Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, 02 1973. ISSN 0033-068X. doi: 10.1143/PTP.49.652. URL https://doi.org/10.1143/PTP.49.652.

[11] Jure Zupan. Introduction to flavour physics. *CERN Yellow Rep. School Proc.*, 6: 181–212, 2019. doi: http://dx.doi.org/10.23730/CYRSP-2019-006.181.

[12] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[13] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972. [Yad. Fiz.15,781(1972)].

[14] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126:298–318, 1977. doi: 10.1016/0550-3213(77)90384-4.

[15] S Camarda. Lhc measurements for constraining parton distribution functions. Saclay, 2017.

[16] The CTEQ Collaboration. The coordinated theoretical-experimental project on qcd, 2019. URL https://www.physics.smu.edu/scalise/cteq/.

[17] The NNPDF Collaboration. Nnpdf, 2019. URL http://nnpdf.mi.infn.it.

[18] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the k_2^0 meson. *Phys. Rev. Lett.*, 13:138–140, Jul 1964. doi: 10.1103/PhysRevLett.13.138. URL https://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[19] David E. Kahana and Sidney H. Kahana. Top and Higgs masses in dynamical symmetry breaking. *Phys. Rev.*, D52:3065–3071, 1995. doi: 10.1103/PhysRevD.52.3065.

[20] F. Abe et al. Observation of top quark production in $\overline{p}p$ collisions with the collider detector at fermilab. *Phys. Rev. Lett.*, 74:2626–2631, Apr 1995. doi: 10.1103/PhysRevLett.74.2626. URL https://link.aps.org/doi/10.1103/PhysRevLett.74.2626.

[21] S. Abachi et al. Observation of the top quark. *Phys. Rev. Lett.*, 74:2632–2637, Apr 1995. doi: 10.1103/PhysRevLett.74.2632. URL https://link.aps.org/doi/10.1103/PhysRevLett.74.2632.

[22] Vittorio del Duca and Eric Laenen. Top physics at the LHC. *Int. J. Mod. Phys.*, A30(35):1530063, 2015. doi: 10.1142/S0217751X1530063X.

[23] Fedor Bezrukov and Mikhail Shaposhnikov. Why should we care about the top quark Yukawa coupling? *J. Exp. Theor. Phys.*, 120:335–343, 2015. doi: 10.1134/S1063776115030152. [Zh. Eksp. Teor. Fiz.147,389(2015)].

[24] Kevin Kroninger, Andreas B. Meyer, and Peter Uwer. Top-Quark Physics at the LHC. In Thomas Schorner-Sadenius, editor, *The Large Hadron Collider: Harvest of Run 1*, pages 259–300. 2015. doi: 10.1007/978-3-319-15001-7_7.

[25] LHC Top Working Group. Lhctopwg summary plots, 2019. URL https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots.

[26] Roel Aaij et al. First observation of top quark production in the forward region. *Phys. Rev. Lett.*, 115(11):112001, 2015. doi: 10.1103/PhysRevLett.115.112001.

[27] Rhorry Gauld. Feasibility of top quark measurements at LHCb and constraints on the large-x gluon PDF. *JHEP*, 02:126, 2014. doi: 10.1007/JHEP02(2014)126.

[28] V. M. Abazov et al. A precision measurement of the mass of the top quark. *Nature*, 429:638–642, 2004. doi: 10.1038/nature02589.

[29] Jean-Luc Caron. Overall view of LHC experiments.. Vue d'ensemble des experiences du LHC. AC Collection. Legacy of AC. Pictures from 1992 to 2002., May 1998. URL https://cds.cern.ch/record/841555.

[30] Esma Mobs. The CERN accelerator complex. Jul 2016. URL https://cds.cern.ch/record/2197559. General Photo.

[31] Daniel Boussard and Trevor Paul R Linnecar. The LHC Superconducting RF System. Technical Report LHC-Project-Report-316. CERN-LHC-Project-Report-316, CERN, Geneva, Dec 1999. URL https://cds.cern.ch/record/410377.

[32] M Bajko et al. Report of the Task Force on the Incident of 19th September 2008 at the LHC. Technical Report LHC-PROJECT-Report-1168. CERN-LHC-PROJECT-Report-1168, CERN, Geneva, Mar 2009. URL https://cds.cern.ch/record/1168025.

[33] Roel Aaij et al. Precision luminosity measurements at LHCb. *JINST*, 9(12):P12005, 2014. doi: 10.1088/1748-0221/9/12/P12005.

[34] F. Follin and D. Jacquet. Implementation and experience with luminosity levelling with offset beam. In *Proceedings, ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders (BB2013): CERN, Geneva, Switzerland, March 18-22 2013*, pages 183–187, 2014. doi: 10.5170/CERN-2014-004.183. [,183(2014)].

[35] R. Jacobsson. Future wishes and constraints from the experiments at the LHC for the Proton-Proton programme. In *Proceedings, ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders (BB2013): CERN, Geneva, Switzerland, March 18-22 2013*, pages 167–176, 2014. doi: 10.5170/CERN-2014-004.167. [,167(2014)].

[36] LHCb Collaboration. 3rd december 2018: End of the first phase and beginning of a new lhcb, 2018. URL http://lhcb-public.web.cern.ch/lhcb-public/.

[37] S. van der Meer. Calibration of the Effective Beam Height in the ISR. 1968.

[38] Colin Barschel. Precision luminosity measurement at LHCb with beam-gas imaging, 2014. URL https://cds.cern.ch/record/1693671. Presented 05 Mar 2014.

[39] Vladislav Balagura. Luminosity at LHCb. Feb 2017. URL https://cds.cern.ch/record/2255091.

[40] Christian Elsässer. $\bar{b}b$ production angle plots. https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[41] The LHCb Collaboration. The LHCb detector at the LHC. *Journal of Instrumentation*, 3(08):S08005–S08005, aug 2008. doi: 10.1088/1748-0221/3/08/s08005. URL https://doi.org/10.1088/1748-0221/3/08/S08005.

[42] P R Barbosa-Marinho et al. *LHCb VELO (VErtex LOcator): Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2001. URL https://cds.cern.ch/record/504321.

[43] S. Viret, C. Parkes, and M. Gersabeck. Alignment procedure of the LHCb Vertex Detector. *Nucl. Instrum. Meth.*, A596:157–163, 2008. doi: 10.1016/j.nima.2008.07.153.

[44] R. Aaij et al. Performance of the LHCb Vertex Locator. *JINST*, 9:P09007, 2014. doi: 10.1088/1748-0221/9/09/P09007.

[45] LHCb magnet: Technical design report. 2000.

[46] Antunes-Nobrega et al. *LHCb reoptimized detector design and performance: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2003. URL https://cds.cern.ch/record/630827.

[47] P R Barbosa-Marinho et al. Technical Design Report LHCb. CERN, Geneva, 2002. URL https://cds.cern.ch/record/582793. revised version number 1 submitted on 2002-11-13 14:14:34.

[48] R Arink et al. Performance of the LHCb Outer Tracker. *JINST*, 9(01):P01002, 2014. doi: 10.1088/1748-0221/9/01/P01002.

[49] M. Adinolfi et al. Performance of the LHCb RICH detector at the LHC. *Eur. Phys. J.*, C73:2431, 2013. doi: 10.1140/epjc/s10052-013-2431-9.

[50] Neville Harnew. Particle identification with the lhcb rich system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, pages 326–332, 07 2006. doi: 10.1016/j.nima.2006.02.167.

[51] Irina Machikhiliyan. First years of running for the LHCb Calorimeter System. Apr 2013. URL https://cds.cern.ch/record/1543420.

[52] P R Barbosa-Marinho et al. *LHCb muon system: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2001. URL https://cds.cern.ch/record/504326.

[53] S Furcas and A Sarti. The LHCb Muon detector commissioning and first running scenarios. *Nucl. Instrum. Methods Phys. Res., A*, 617(LHCb-PROC-2009-025. LHCb-CONF-2009-025. CERN-LHCb-CONF-2009-025. 1-3):158–160. 3 p, Aug 2009. doi: 10.1016/j.nima.2009.06.071. URL https://cds.cern.ch/record/1205065.

[54] E Rodrigues. Tracking definitions. Technical Report LHCb-2007-006. CERN-LHCb-2007-006, CERN, Geneva, Feb 2007. URL https://cds.cern.ch/record/1016937. revised version submitted on 2007-03-28 09:34:37.

[55] LHCb tracking group. Tracking and alignment plots for conferences. URL https://twiki.cern.ch/twiki/bin/view/LHCb/ConferencePlots.

[56] LHCb Tracking Group. Pat seeding. URL https://twiki.cern.ch/twiki/bin/view/LHCb/PatSeeding.

[57] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.*, A262:444–450, 1987. doi: 10.1016/0168-9002(87)90887-4.

[58] Roel Aaij et al. Measurement of the track reconstruction efficiency at LHCb. *JINST*, 10(02):P02007, 2015. doi: 10.1088/1748-0221/10/02/P02007.

[59] Roel Aaij et al. LHCb Detector Performance. *Int. J. Mod. Phys.*, A30(07):1530022, 2015. doi: 10.1142/S0217751X15300227.

[60] Marcin Kucharczyk, Piotr Morawski, and Mariusz Witek. Primary Vertex Reconstruction at LHCb. Technical Report LHCb-PUB-2014-044. CERN-LHCb-PUB-2014-044, CERN, Geneva, Sep 2014. URL https://cds.cern.ch/record/1756296.

[61] F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel. *Robust statistics: the approach based on influence functions*. Wiley series in probability and mathematical statistics. Probability and mathematical statistics. Wiley, 2005. ISBN 9780471735779. URL https://books.google.co.uk/books?id=NhbvAAAAMAAJ.

[62] Christian Lippmann. Particle identification. *Nucl. Instrum. Meth.*, A666:148–172, 2012. doi: 10.1016/j.nima.2011.03.009.

[63] Roel Aaij, Johannes Albrecht, Francesco Dettori, Kevin Dungs, Helder Lopes, Diego Martinez Santos, Jessica Prisciandaro, Barbara Sciascia, Vasileios Syropoulos, Sascha Stahl, and Ricardo Vazquez Gomez. Optimization of the muon reconstruction algorithms for LHCb Run 2. Technical Report LHCb-PUB-2017-007. CERN-LHCb-PUB-2017-007, CERN, Geneva, Feb 2017. URL http://cds.cern.ch/record/2253050.

[64] William Barter, Albert Bursche, Victor Coco, Oscar Francisco, CÃľdric Potterat, and Murilo Rangel. Jets reconstruction and performances at LHCb. Technical Report LHCb-INT-2012-015. CERN-LHCb-INT-2012-015, CERN, Geneva, May 2012. URL https://cds.cern.ch/record/1446552.

[65] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008. doi: 10.1088/1126-6708/2008/04/063.

[66] O A De Aguiar Francisco, M Potterat, and M Rangel. LHCb jet identification. Technical Report LHCb-INT-2014-021. CERN-LHCb-INT-2014-021, CERN, Geneva, Mar 2014. URL https://cds.cern.ch/record/1690574.

[67] R Antunes-Nobrega et al. *LHCb trigger system: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2003. URL http://cds.cern.ch/record/630828. revised version number 1 submitted on 2003-09-24 12:12:22.

[68] LHCb trigger group. Trigger schemes. URL http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html.

[69] T Head. The LHCb trigger system. *Journal of Instrumentation*, 9(09):C09015–C09015, sep 2014. doi: 10.1088/1748-0221/9/09/c09015. URL https://doi.org/10.1088%2F1748-0221%2F9%2F09%2Fc09015.

[70] Barbara Sciascia. LHCb Run 2 Trigger Performance. *PoS*, BEAUTY2016(LHCb-PROC-2016-020. CERN-LHCb-PROC-2016-020):029. 7, May 2016. doi: 10.22323/1.273.0029. URL https://cds.cern.ch/record/2208038.

[71] Giulio Dujany and Barbara Storaci. Real-time alignment and calibration of the LHCb Detector in Run II. *J. Phys.: Conf. Ser.*, 664(LHCb-PROC-2015-011. CERN-LHCb-PROC-2015-011):082010. 8 p, Apr 2015. doi: 10.1088/1742-6596/664/8/082010. URL https://cds.cern.ch/record/2017839.

[72] G. Barrand et al. GAUDI - A software architecture and framework for building HEP data processing applications. *Comput. Phys. Commun.*, 140:45–55, 2001. doi: 10.1016/S0010-4655(01)00254-5.

[73] Lund University. Introduction to pythia. URL http://home.thep.lu.se/~torbjorn/Pythia.html.

[74] I. Belyaev et al. Simulation application for the LHCb experiment,. *eConf*, C0303241:TUMT003, 2013.

[75] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506: 250–303, 2003. doi: 10.1016/S0168-9002(03)01368-8.

[76] LHCb Collaboration. The boole project, . URL http://lhcbdoc.web.cern.ch/lhcbdoc/boole/.

[77] LHCb Collaboration. The brunel project, . URL http://lhcbdoc.web.cern.ch/lhcbdoc/brunel/.

[78] LHCb Collaboration. The moore project, . URL http://lhcbdoc.web.cern.ch/lhcbdoc/moore.

[79] D Britton and S L Lloyd. How to deal with petabytes of data: the LHC grid project. *Reports on Progress in Physics*, 77(6):065902, jun 2014. doi: 10.1088/0034-4885/77/6/065902. URL https://doi.org/10.1088%2F0034-4885%2F77%2F6%2F065902.

[80] F Stagni et al. LHCbDirac: distributed computing in LHCb. *Journal of Physics: Conference Series*, 396(3):032104, dec 2012. doi: 10.1088/1742-6596/396/3/032104. URL https://doi.org/10.1088%2F1742-6596%2F396%2F3%2F032104.

[81] Frederic Brochu et al. Ganga: a tool for computational-task management and easy access to grid resources. *CoRR*, abs/0902.2685, 2009. URL http://arxiv.org/abs/0902.2685.

[82] LHCb Collaboration. The davinci project, . URL http://lhcbdoc.web.cern.ch/lhcbdoc/davinci/.

[83] Henry Brown. Monitoring Radiation Damage in the Vertex Locator and Top Pair Production in LHCb, Sep 2013. URL http://cds.cern.ch/record/1635683. Presented 07 Nov 2013.

[84] LHCb Collaboration. Electroweak physics: Useful plots, . URL http://twiki.cern.ch/twiki/bin/view/LHCbPhysics/UsefulPlots.

[85] Rhorry Gauld. Measuring top quark production asymmetries at LHCb. Technical Report LHCb-PUB-2013-009. CERN-LHCb-PUB-2013-009, CERN, Geneva, Jun 2013. URL https://cds.cern.ch/record/1557385.

[86] Roel Aaij et al. Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era. (LHCB-PUB-2018-009. LHCC-G-171), Aug 2018. URL https://cds.cern.ch/record/2636441.

[87] Serguei Chatrchyan et al. Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 02:024, 2014. doi: 10.1007/JHEP02(2014)024,10.1007/JHEP02(2014)102. [Erratum: JHEP02,102(2014)].

[88] Hung-Liang Lai et al. Parton Distributions for Event Generators. *JHEP*, 04:035, 2010. doi: 10.1007/JHEP04(2010)035.

[89] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. doi: 10.1007/JHEP07(2014)079.

[90] Carlo Oleari. The powheg-box. *Nucl. Phys. Proc. Suppl.*, 205-206:36–41, 2010. doi: 10.1016/j.nuclphysbps.2010.08.016.

[91] Tomas Jezo. Top quark modelling in POWHEG BOX. In *10th International Workshop on Top Quark Physics (TOP2017) Braga, Portugal, September 17-22, 2017*, 2018.

[92] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015. doi: 10.1007/JHEP04(2015)040.

[93] Stephen Farry and William James Barter. Measurement of forward top quark pair production in the μeb final state. *LHCb Analysis Note*, Aug 2017. URL https://cds.cern.ch/record/2276788.

[94] R Aaij et al. Inclusive W and Z production in the forward region at $\sqrt{s} = 7$ TeV. *JHEP*, 06:058, 2012. doi: 10.1007/JHEP06(2012)058.

[95] R Aaij et al. Measurement of the cross-section for $Z \to e^+e^-$ production in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 02:106, 2013. doi: 10.1007/JHEP02(2013)106.

[96] Roel Aaij et al. Identification of beauty and charm quark jets at LHCb. *JINST*, 10(06):P06013, 2015. doi: 10.1088/1748-0221/10/06/P06013.

[97] Emanuele Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *Eur. Phys. J.*, C71:1547, 2011. doi: 10.1140/epjc/s10052-011-1547-z.

[98] S Farry and N Chiapolini. A measurement of high-pT muon reconstruction efficiencies in 2011 and 2012 data. Technical Report LHCb-INT-2014-030. CERN-LHCb-INT-2014-030, CERN, Geneva, Jul 2014. URL https://cds.cern.ch/record/1709688.

[99] David Ward. Measurement of $Z \to e^+e^-$ production at $\sqrt(s) = 13$TeV. Apr 2016. URL https://cds.cern.ch/record/2147394.

[100] Roel Aaij et al. Measurement of forward W and Z boson production in association with jets in proton-proton collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 05:131, 2016. doi: 10.1007/JHEP05(2016)131.

[101] S. Dittmaier et al. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. 2011. doi: 10.5170/CERN-2011-002.

[102] S. Schael et al. Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP. *Phys. Rept.*, 532:119–244, 2013. doi: 10.1016/j.physrep.2013.07.004.

[103] F. Abe et al. Evidence for w^+w^- production in $\overline{p}p$ collisions at $\sqrt{s} = 1.8hrmTeV$. *Phys. Rev. Lett.*, 78:4536–4540, Jun 1997. doi: 10.1103/PhysRevLett.78.4536. URL https://link.aps.org/doi/10.1103/PhysRevLett.78.4536.

[104] V. M. Abazov et al. Measurement of the WW production cross section with dilepton final states in p anti-p collisions at $\sqrt{s} = 1.96$ TeV and limits on anomalous trilinear gauge couplings. *Phys. Rev. Lett.*, 103:191801, 2009. doi: 10.1103/PhysRevLett.103.191801.

[105] Georges Aad et al. Measurement of W^+W^- production in pp collisions at $\sqrt{s}$=7 TeV with the ATLAS detector and limits on anomalous WWZ and WWγ couplings.

Phys. Rev., D87(11):112001, 2013. doi: 10.1103/PhysRevD.87.112001,10.1103/PhysRevD.88.079906. [Erratum: Phys. Rev.D88,no.7,079906(2013)].

[106] Georges Aad et al. Measurement of total and differential W^+W^- production cross sections in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector and limits on anomalous triple-gauge-boson couplings. *JHEP*, 09:029, 2016. doi: 10.1007/JHEP09(2016)029.

[107] Morad Aaboud et al. Measurement of the W^+W^- production cross section in pp collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV with the ATLAS experiment. *Phys. Lett.*, B773:354–374, 2017. doi: 10.1016/j.physletb.2017.08.047.

[108] Morad Aaboud et al. Measurement of fiducial and differential W^+W^- production cross-sections at $\sqrt{s}$ =13 TeV with the ATLAS detector. 2019.

[109] Serguei Chatrchyan et al. Measurement of the W^+W^- Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV and Limits on Anomalous $WW\gamma$ and WWZ Couplings. *Eur. Phys. J.*, C73(10):2610, 2013. doi: 10.1140/epjc/s10052-013-2610-8.

[110] Vardan Khachatryan et al. Measurement of the W^+W^- cross section in pp collisions at $\sqrt{s} = 8$ TeV and limits on anomalous gauge couplings. *Eur. Phys. J.*, C76 (7):401, 2016. doi: 10.1140/epjc/s10052-016-4219-1.

[111] Measurement of the WW cross section pp collisions at $\sqrt{s}$=13 TeV. Technical Report CMS-PAS-SMP-16-006, CERN, Geneva, 2016. URL https://cds.cern.ch/record/2160868.

[112] CMS Collaboration. Summaries of cms cross section measurements, 2019. URL http://cern.ch/go/pNj7.

[113] Roel Aaij et al. Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 09:136, 2016. doi: 10.1007/JHEP09(2016)136.